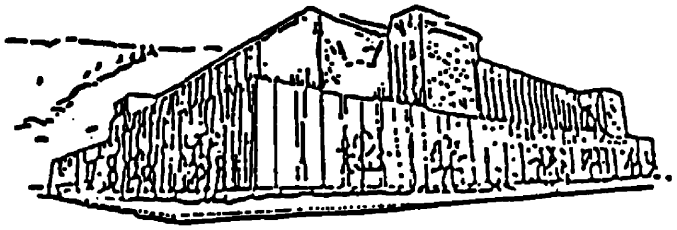
1999

Effect of DNA polymerases on the production of short tandem repeat stutter bands

James M. Streeter
The University of Montana



Maureen and Mike
MANSFIELD LIBRARY

The University of **MONTANA**

** *Please check "Yes" or "No" and provide signature* **

Yes, I grant permission ✓
No, I do not grant permission ____

Author's Signature James M. Streeter

Date July 17, 1999

THE EFFECT OF DNA POLYMERASES ON THE PRODUCTION OF SHORT TANDEM REPEAT STUTTER BANDS

BY

JAMES M. STREETER

B.A., CALIFORNIA STATE UNIVERSITY AT SACRAMENTO, 1973

PRESENTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

MASTER OF SCIENCE

THE UNIVERSITY OF MONTANA

1999

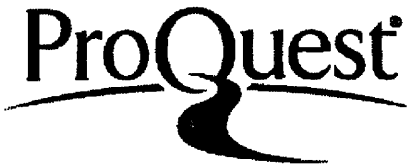
Approved By:

Chairperson

Dean, Graduate School

7/16/99

Date



ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 - 1346

Streeter, James M. July 1999 **Microbiology**

The Effect of DNA Polymerases on the Production of Short Tandem Repeat Stutter Bands

Director: D. Scott Samuels

Abstract:

PCR (polymerase chain reaction) has revolutionized the practice of molecular forensics. The problems of degraded DNA, small sample quantity, and DNA mixtures make forensic materials unique. Using PCR as a tool for the preparation of forensic samples, multiple DNA markers can be readily identified in very small and degraded samples. Amplifying several DNA markers from the same starting template increases the power of exclusion and allows for the identification of mixtures.

Recent advances in the analysis of short segments of tandemly repeated DNA, called short tandem repeats (STRs), has allowed for the personal identification of donors of DNA from forensic samples. One inherent problem with STRs is the production of PCR products called stutter bands. The presence of these bands can make the interpretation of results difficult. Since crimes usually involved at least two individuals, the possibility exists for a mixture of DNAs in a sample.

My research focused on the possible elimination or lessening of these unwanted PCR products using the STR multiplex called PowerPlex™. The Promega Corporation PowerPlex™ simultaneously coamplifies eight STRs: CSF1PO, TPOX, THO1, vWA, D16S539, D7S820, D13S317, and D5S818. In my research, the conditions for the PCR reaction, the type and manufacturer of the DNA polymerase enzyme and the amount of template were changed. These variations resulted in little or no reduction of stutter bands. A survey of samples was also examined to determine the range of stutter bands present at each locus. In conclusion, the presence of stutter bands does not appear to be related to the type, concentration, or manufacturer of the DNA polymerase but appears to be a consequence of the STRs internal repeat sequence.

Table of Contents

List of Tables

List of Figures

Chapter 1-Introduction

Personal identification of the source of physical evidence has always been the goal of the forensic scientists. Fingerprints are the most common method of personal identification in forensic samples. Over a hundred years ago in England, the "Troup Committee", searching for the best method for identifying habitual criminals, stated that the chance of two fingerprints being identical is less than one in sixty-four trillion (Browne, 1954). In the thirty-first chapter of *Mark Twain's Life on the Mississippi*, "A Thumbprint and What Came of It," Mark Twain wrote one of the first detective stories that dealt with fingerprint evidence.

Forensic serology is the analysis of biological samples. These biological samples may be blood, semen or saliva that has been transferred from one individual to another or from one individual to an object, such as a car seat or weapon. Only since the development of DNA markers has the goal of individualization of biological stains become a reality.

Blood types A, B, AB, and O were the first recognized difference in human blood. These markers, called the ABO blood system, were discovered in 1901 by Dr. Karl Landsteiner (Landsteiner, 1901). Use of these genetic markers in the forensic laboratory was handicapped by the potential for degradation of the sample over time or by environmental conditions. In addition, only four groups could be classified, so the degree of discrimination was limited (ABO type O $\approx$ 46% of the general population), and typing was subject to interference from substrates. Finally, the total amount of blood needed to perform a test was significant.

The next landmark in forensic serology was electrophoretic separation of polymorphic proteins and enzymes. These genetic markers allowed the forensic samples to be

separated into additional groups (Culliford, 1967; Wraxall, 1968), thus increasing power of discrimination. Table 1 shows several polymorphic enzymes and the number of associated variants.

Table 1. Examples of Forensic Polymorphic Enzymes

Enzyme Name	Initials	Number of Common Variants
Red Cell Acid Phosphatase	EAP	6
Phosphoglucomutase	PGM	10
Adenylate Kinase	AK	3

Number of common variants detected in three genetic markers that were used by forensic science laboratories. Each conventional enzyme must be analyzed separately and required an additional amount of biological sample.

If the genetic markers are independently inherited, then the probability that they will occur together is the product of the each individual probability and is referred to as the "Product Rule" (Saferstein, 1988). Table 2 is an example of the product rule applied to five conventional genetic markers. These probabilities are contrasted with the probabilities determined by STR markers

Table 2. Comparison of the Power of Exclusion (Conventional vs. DNA)

Blood Group System	Phenotype Frequency (Whites)
ABO Type O	0.46
AK 1	0.95
PGM 2+1+	0.23
Es-D 2-1	0.19
EAP B	0.39
Total Probability all five conventional markers occurring together	0.013, or 1/79
PowerPlex 1.1 System-8 STR markers. Total Probability all eight DNA markers occurring together	0.000000043496, or 1/2,299,000,000

Difference in the power to discriminate individuals by analyzing the same sample with five conventional markers or eight STRs. By using the product rule, the power of discrimination can be greater increased by analyzing the sample with eight STRs instead of five conventional markers.

"DNA fingerprinting" forever changed forensic serology. Genetic variation at a single locus using the technique called RFLP (restriction fragment length polymorphism) was first described in the literature in 1980 (Wyman, 1980). In this technique, DNA is cut at specific sites by restriction enzymes. The generated fragments are separated on an agarose gel by their size. Variations in the sizes of the fragments, VNTR (variable number of tandem repeats), were identified by use of radioactive probes and visualized on x-ray film. The pattern on the x-ray film is similar to a series of bar codes. Alec Jeffreys was the first scientist to successfully use this technique in criminal cases. He used it first to exclude a potential suspect and eventually to identify the assailant in two related rape/homicides. The technique was termed "DNA fingerprinting" (Jeffreys, 1985; Jeffreys, 1986). RFLP suffers from the limitation of the quantity of high molecular weight DNA needed for the analysis, gel variations, intrinsic measurement variability of the allele sizing and time required for performing the analysis.

The isolation and identification of a thermally stable DNA polymerase changed the analysis of nucleic acids. Use of this enzyme allowed for the automated amplification of selected areas of DNA. This amplification process is called PCR (polymerase chain reaction) and its the potential for use in forensics was described in the late 1980's (Erlich, 1989; Erlich, 1991; Gibbs, 1990; Reynolds, 1991).

The original adaptation of PCR in forensics identified sequence-based variations (Comey, 1991). Samples were typed with allele-specific oligonucleotide probes using the reverse dot blot procedure. The probes were attached to specific areas on a plastic strip. The DNA was hybridized to its corresponding probe and visualized using a colorimetric assay. Difficulties with this technique included the low number of variants and interpretation of mixtures because of the color intensity of the dots.

PCR products are separated by length electrophoretically. The first validated length-based PCR marker was an AFLP (amplified fragment length polymorphism) called D1S80 (Cosso, 1995). The amplified D1S80 fragments are electrophoretically separated, chemically stained, and compared to allelic ladders run in adjacent lanes. The PCR product size varies from allele to allele depending on the number of 16 base pair repeats. Ninety percent of the alleles are 14 to 41 repeat units. Because of preferential amplification of the short alleles over the longer alleles, degraded samples are difficult to interpret.

The current method of choice for forensic DNA typing is the characterization of STRs (short tandem repeats) (Edwards, 1991; Edwards, 1992) after amplification by PCR (Reynolds, 1991). Since forensic samples are often degraded, VNTRs are not processed by PCR due to their large size.

STRs are tandemly repeated two to seven base pair sequence motifs. The number of copies of the repeat sequence differentiates the STR alleles. As an example, if four copies of the repeat sequence are present, then the STR has allele four. Over 3600 polymorphic STRs have been identified in the human genome (Murray, 1994).

After electrophoretic separation, the alleles are identified using radioactive probes, silver stain, or fluorescence (Buel, 1998; Worley, 1994). Since the STR products are of discrete and separable lengths, they can be identified by comparison to allelic ladders resolved on the electrophoretic gel (Puers, 1993).

Their size variations (120-330 bp) (Fregeau, 1993; Kline, 1997; Sprecher, 1996) allow the amplification of several repeat sequences in the same reaction mixture, which is called multiplexing (Lins, 1996; Ricciardone, 1997; Schumm, 1997). Multiplex STR kits have been validated for use in forensic laboratories (Budowle, 1997; Kline, 1997; Micka, 1996).

One PCR artifact of STR amplification is minor bands that appear one repeat band below the true allele. These minor bands are called stutter bands, DNA polymerase slippage products, or shadow bands. Sequence analysis of stutter bands show that they lack one repeat unit relative to the main allele (Murray, 1993). Forensic samples often contain a mixture of DNA from several individuals, blood on blood, blood on semen, etc. The presence of stutter bands can confuse the interpretation of the DNA profiles.

Fig. 1 shows the results from several samples typed with the four STR markers CSF1PO, TPOX, THO1, and vWA, and Fig. 2 shows the same samples typed with the STR markers D16S539, D7S820, D13S317, and D5S818. The DNA samples involve an alleged sexual assault. Two suspects are alleged to have sexually assaulted a female

victim and the samples show the results from the victim's underwear and sperm fraction of the vaginal swab. Samples from the victim's underwear are found in two lanes of each gel because the DNA extracted from that item is divided into two portions, the Epi or female fraction and the Sperm fraction from the male(s).

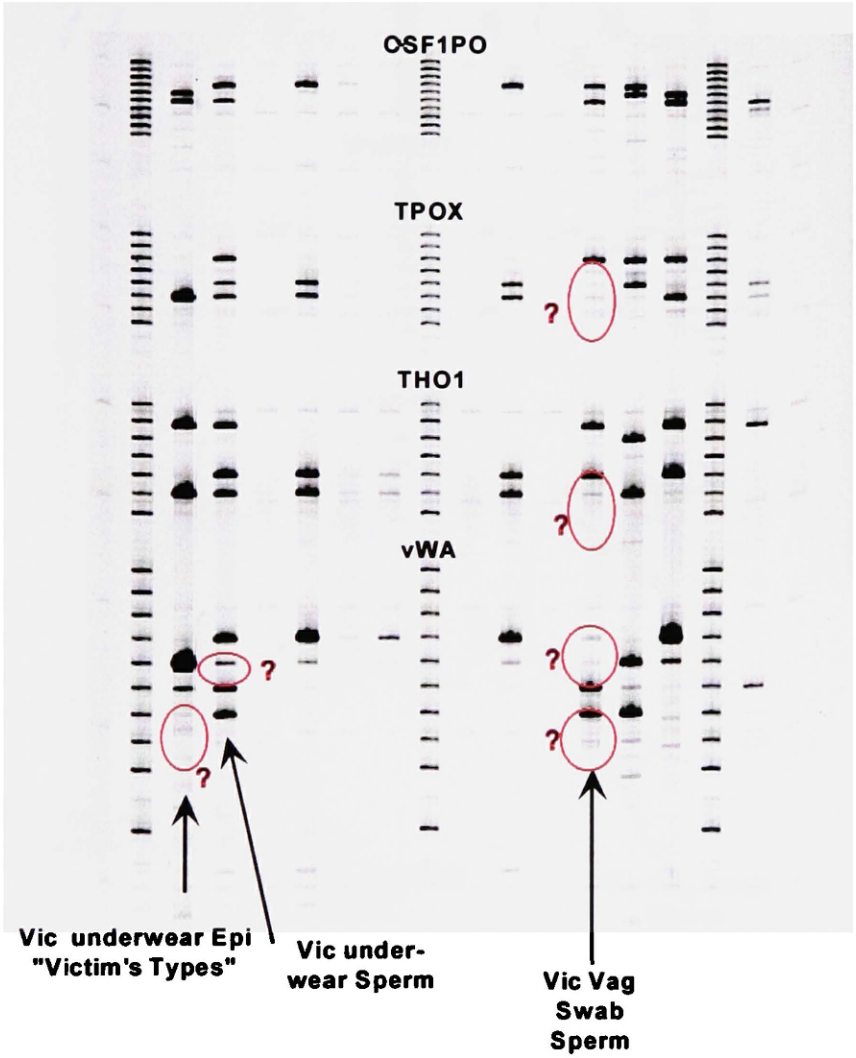


Fig. 1. An electrophoretic gel showing the results of four STR markers CSF1PO, TPOX, THO1 and vWA. The samples are applied to the gel at the top and separated by fragment lengths. Three of the samples applied to the gel are samples from a female victim of a sexual assault. The commonly found variants in the general population are both end rows and in the middle row. The dark bands are the true STR variants. The red circles show the presence of additional DNA or PCR artifacts, including stutter bands. The samples were amplified with the Promega Corporation PowerPlex™ kit (Riley, 1997; Schumm, 1997).

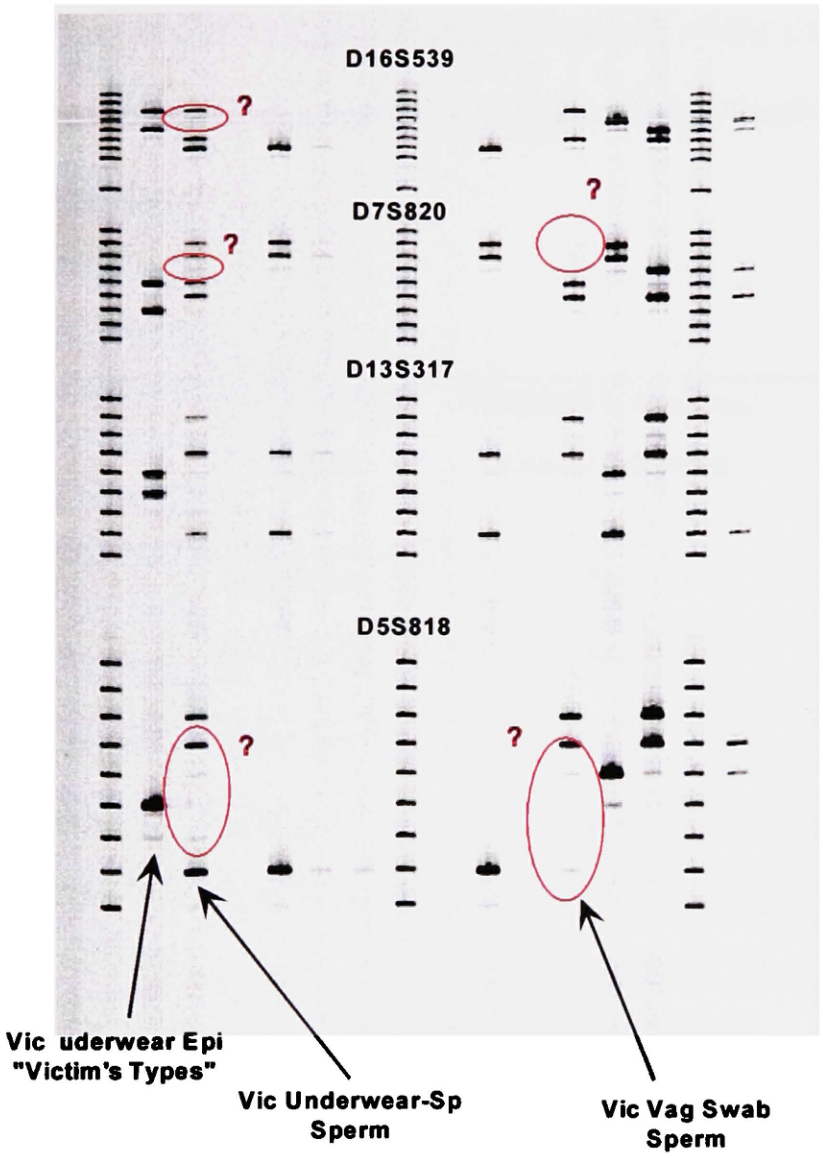


Fig. 2. An electrophoretic gel showing the results of four STR markers D16S539, D7S820, D13S317, and D5S818. These are the same samples as seen in Fig. 1. The commonly found variants in the general population are both end rows and in the middle row. The dark bands are the true STR variants. The red circles show the presence of additional DNA or PCR artifacts, including stutter bands. The samples were amplified with the Promega Corporation PowerPlex™ kit (Riley, 1997; Schumm, 1997).

Clearly the results demonstrate the presence of DNA from more than one individual and faint bands from other DNA donors or stutter bands. The possible presence of stutter combined with multiple DNA donors make interpretation of the results difficult.

In the PowerPlex™ system, eight STRs, CSF1PO, TPOX, THO1, vWA, D16S539, D7S820, D13S317, and D5S818, are amplified simultaneously. Table 3 shows the PowerPlex™ STRs, their fluorescent label, chromosomal location, GenBank® designation, and repeat sequence.

Table 3-PowerPlex™ Loci Information

STR Locus	Fluorescent Label	Chromosomal Location	GenBank® locus and Locus Definition	Repeat Sequence (ISFH)
D16S539	Fluorescein	16q24-qter	NA	GATA
D7S820	Fluorescein	7q11.21-22	NA	GATA
D13S317	Fluorescein	13q22-q31	NA	TATC
D5S818	Fluorescein	5q23.3-32	NA	AGAT
CSF1PO	TMR	5q33.3-34	HUMCSF1PO, human c-fms proto-oncogene for CSF-1 receptor gene	TAGA
TPOX	TMR	2p25.1-pter	HUMTPOX, human thyroid peroxidase gene	TGAA
THO1	TMR	11p15.5	HUMTHO1, human tyrosine hydroxylase gene	TCAT
vWA	TMR	12p12-pter	HUMVWFA31, human von Willebrand factor gene	TCTA

Locus information for the eight STR markers amplified in the PowerPlex™ 1.1 system (Lins, 1998). The CTTv markers are labeled with fluorescein. The DDDD markers are labeled with TMR (carboxy-tetramethylrhodamine).
ISFH = International Society for Forensic Haemogenetics

During the PCR one strand of DNA is labeled with a fluorescently tagged primer. The fluorescently tagged STR bands are visualized after electrophoresis in the FMBIO™ II

fluorescent imager (Hitachi, 1998) as two sets of four markers. The allele type or allele designation in each sample can be identified by visual comparison to allelic ladders run on the same gel or identified using a computer software package called STaR™ call. Alleles of STR loci are differentiated by the number of copies of the repeat sequence contained within the amplified region. The STaR™ call software will identify the specific allele by size (base pair) and allele intensity (OD-optical density) as measured by peak height.

The fidelity of the DNA polymerase to accurately produce a daughter strand has an influence upon the production of stutter bands. Errors in DNA polymerase fidelity are divided into two classes. One class of errors is the substitution of incorrect nucleotides during DNA synthesis and is called polymerase error rate. The other class of errors is when extra nucleotides are inserted or nucleotides are left out during synthesis, called frameshift errors. Frameshift errors are affected by the structure and processivity of the DNA polymerase, length and base composition of the sequence to be amplified, the influence of neighboring bases (Bebenek, 1993; Lewin, 1997), and template-primer misalignment (Eckert, 1993).

The DNA polymerase will dissociate from the template-primer after the addition of a certain number of nucleotides. The processivity value can be measured as the number of incorporated nucleotides before dissociation.

Taking into account the processivity of DNA polymerase, a slipped strand mispairing model has been proposed to explain stutter band production during amplification of dinucleotides repeats (Levinson, 1987; Schlotterer, 1992; Thacker, 1992) and frameshifts on short sequences (Kunkel, 1988). In the slipped strand model shown in Fig. 3, the

DNA polymerase dissociates from the template strand and the partially synthesized daughter strand will denature from the template. A single repeat unit can then loop out in the template strand before the two strands re-associate. When the synthesis is completed, the new strand will have one less repeat unit than the template strand.

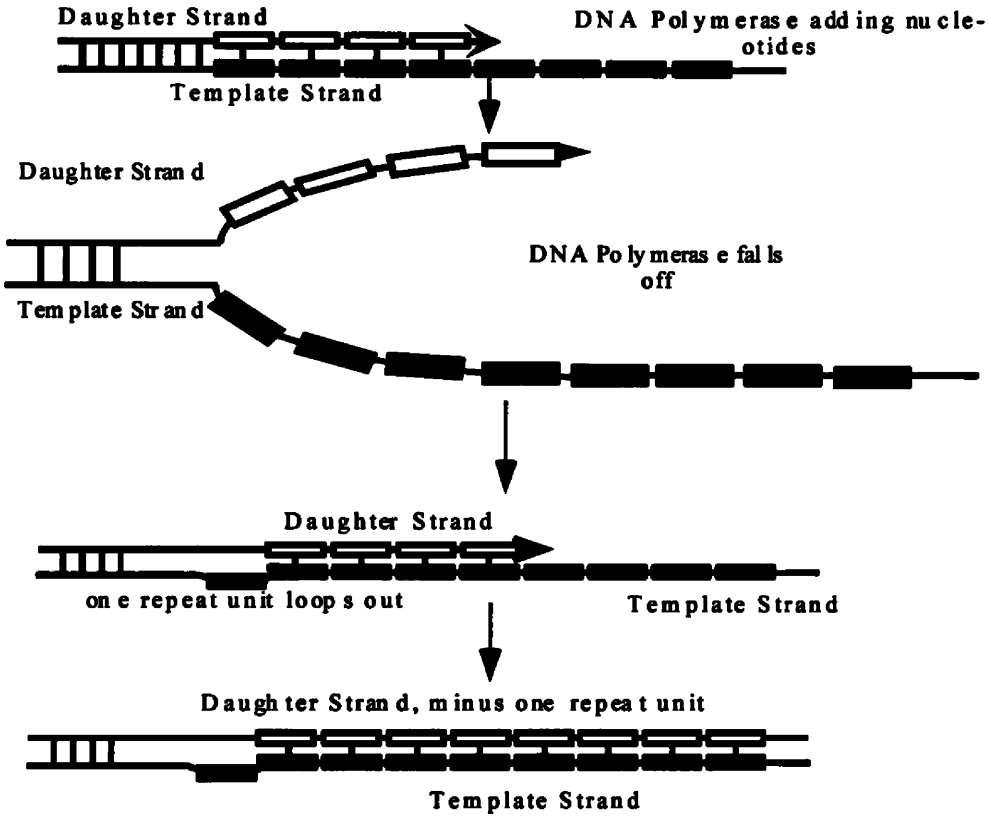


Fig. 3. Diagrammatic representation of the slipped strand synthesis model. In the first step the DNA polymerase is adding nucleotides to the daughter strand. The DNA polymerase will dissociate and two DNA strands can separate, this is shown in the second step. If the two strands come together and the number of repeats is not correctly aligned, as seen in the third step, the daughter stand can be synthesized with one less repeat unit, as seen in step four.

Strand slippage can also explain DNA polymerase-mediated deletions of hundreds of nucleotides between direct repeat sequences (Kunkel, 1992). Sequence analysis of stutter bands have shown that they lack one repeat unit (Murray, 1993). These scientists also

suggested that a more processive thermostable DNA polymerase could result in the reduction (or elimination) of the stutter bands.

Recent work on the STR HUMFibra/FGA has shown that substituting different thermostable DNA polymerases that have less processivity than *Taq* resulted in an increase in the amount of stutter (Table 4) (Meldguard, 1998). However, some researchers have reported that processivity is not an important determinant in eucaryotic replication fidelity (Abbotts, 1984).

Table 4. DNA Polymerase Processivity

DNA Polymerase	Processivity (nucleotides)	% stutter (HUMFibra/FGA)
Taq	50-60	7%
Pfu	10-15	13%
Pwo	20-30	17%
Vent	10	14%

Difference in processivity of four DNA polymerases and the observed differences in percentage stutter with the STR HUMFibra/FGA (Meldguard, 1998). Processivity is the number of nucleotides that are incorporated before the DNA polymerase dissociates.

Other conditions that affect DNA polymerase fidelity include temperature, dNTP concentration, salt concentrations, and pH (Eckert, 1993; Keohavong, 1989); however optimizing these parameters did not significantly affect the proportion of dinucleotide amplification frameshift products (Hite, 1996).

I investigated the theory that by changing the thermal cycle conditions, type of DNA polymerase, amount of DNA polymerase, or amount of template, I could reduce or eliminate the production of stutter bands in the PowerPlex™ system.

Chapter 2-Methods

The following procedures were followed for each test sample:

1) Samples were collected from individuals in the form of blood samples or bucal swabs (mouth swabs). The DNA was isolated and purified by a phenol/chloroform extraction and Microcon™ concentration or Chelex® extraction.

2) The amount of extracted DNA in each sample was measured by using a Quantiblot™ kit, a human-specific slot blot quantitation kit.

3) The samples were amplified by PCR with the PowerPlex™ system. The kit contains the STR buffer with dNTPs, and 16 primers to amplify all eight STRs simultaneously. The samples that were amplified with different DNA polymerases are noted in the Results section. The DNA polymerase was not included in the commercial kit and was added separately. Table 5 shows the components of a typical PCR master mix.

Table 5. PowerPlex™ Master Mix

PCR Master Mix Components	Volume per Sample (ul)
Sterile Water	18.00
Gold ST*R 10X Reaction Buffer (containing dNTPs, BSA and buffer, pH 8.3)	2.5
Locus-specific primer pairs (8)	2.5
Ampli*Taq* Gold™	1 (1 unit)
Total Volume	24

The components of the PowerPlex™ Master Mix (Promega Corporation, 1997) used in this experiment. One microliter of template was added to each master mix tube to have a final volume of twenty-five microliters. The typical PowerPlex™ master mix uses the Ampli*Taq* Gold™ DNA polymerase. If a different DNA polymerase was used in the experiment that required a reaction buffer of pH 9, STR 10X buffer was substituted.

One ul of sample (1-3 ng of DNA) was added to each reaction mix for a total volume of 25 ul.

4) All samples were amplified in a Perkin Elmer model 480 thermal cycler.

Thermal Cycling Profile: (Promega Corporation, 1997)

95°C incubate for 11 minutes then:

96°C incubate for 2 minutes, then:

94°C for 1 minute

60°C for 1 minute

70°C for 1.5 minutes

For 10 cycles, then:

90°C for 1 minute

60°C for 1 minute

70°C for 1.5 minutes

For 20 cycles, then:

60°C soak for 30 minutes

The initial 95°C incubation is used to activate the Ampli*Taq* Gold™ DNA polymerase and may not be used when other DNA polymerases are substituted.

5) The samples were separated on a denaturing 6% 0.4-mm thick acrylamide gel. After electrophoresis, the gels were placed into the FMBIO™ II for analysis. The instrument contains four filters. Two of the filters are used to analyze four STR markers each and one filter will be used to identify the molecular size marker include in each gel lane. The other filter is used for focusing the instrument. After scanning, each gel was analyzed, the lanes identified, and lane width, position of bands and peak area of each band determined.

6) The peak height of the STR bands, true allele bands and stutter products, in all eight STRs was calculated using the computer software STaR™. The FMBIO™ II instrument records the amount of observed DNA in each STR band as "OD counts". The percent

stutter was calculated as the ratio of the stutter band OD value and the true allele band OD value.

Chapter 3-Results

Population Study

One hundred and fifty one samples were amplified and typed using the PowerPlex™ system. Each sample was examined for the presence of stutter in all eight loci and all observed stutter levels were recorded. Not all of the samples produced stutter bands and not all of the alleles were found in the population survey. The alleles that were found in this study are displayed in Table 6.

Table 6. STR Alleles Observed in This Study

STR	Alleles Found in General Population	Alleles Found in this Study
CSF1PO	6-15	8-14
TPOX	6-13	7-11
THO1	5-11	5-9
vWA	11,13-21	13-19
D16S539	5,8-15	8-13
D7S820	6-14	6-12
D13S317	7-15	7-13
D5S818	7-15	8-12

Comparison of the commonly detected alleles for all eight STRs in the PowerPlex™ 1.1 system with the alleles detected in this study.

With the exception of THO1, as the allele number increased at each locus, the average stutter level also increased. THO1 is the only STR that has a common allele with a repeat sequence that is less than 4 base pair. One of the most common alleles at this locus is allele 9.3. This allele has 3 base pairs more than the 9 allele and 1 base pair less than the 10 allele. Statistical analysis of the stutter levels for the population studies are found in Table 7. This table shows the number of observed samples, and minimum, maximum, and mean of the observed percent stutter for the population study.

Table 7. Distribution of Stutter Among STR Marker Alleles

Marker	Allele	N (number of samples)	Minimum (% Stutter)	Maximum (% Stutter)	Mean (% Stutter)	Std. Deviation
CSF1P0	A8	5	2.2	2.7	2.41	.20
	A9	35	.4	7.3	2.41	1.29
	A10	44	.8	11.0	3.27	1.65
	A11	42	.7	7.4	3.55	1.71
	A12	15	.8	7.5	4.10	1.80
	A13	3	3.4	5.3	4.35	.95
	A14	3	3.9	5.7	4.74	.90
D13S317	A7	7	.40	1.62	.86	.43
	A8	10	.5	2.6	1.40	.66
	A9	6	1.3	3.8	2.65	.87
	A10	12	1.7	6.7	2.90	1.51
	A11	35	.6	10.4	4.44	2.13
	A12	21	.9	8.9	5.03	2.07
	A13	11	3.6	11.1	6.59	2.25
D16S539	A7	2	2.06	2.91	2.49	.60
	A8	18	.4	7.4	3.54	1.77
	A9	19	1.5	6.1	3.62	.92
	A10	54	1.4	8.0	4.12	1.51
	A11	49	.0	13.3	4.47	2.45
	A12	19	.6	14.3	4.79	3.08
	A13	3	3.0	7.0	5.05	2.04
D5S818	A8	4	1.8	4.1	3.02	.93
	A9	17	1.7	7.6	4.04	1.47
	A10	75	.3	11.0	4.88	2.17
	A11	46	2.3	10.0	5.88	1.70
	A12	23	1.2	13.4	7.42	2.57
D7S820	A6	3	1.47	2.70	1.96	.65
	A7	25	.30	9.55	2.44	1.76
	A8	20	.1	6.1	3.02	1.49
	A9	49	1.4	7.1	4.36	1.44
	A10	14	3.0	8.3	5.43	1.73
	A11	26	1.9	8.8	5.46	1.61
	A12	8	3.0	13.1	7.20	3.33
TH01	A5	41	.1	3.2	1.34	.85
	A6	30	.31	6.70	2.72	1.52
	A7	13	1.26	3.99	2.53	.74
	A8	26	.5	6.9	3.31	1.71
	A9	20	.9	6.8	1.96	1.42
TP0X	A7	47	.10	3.20	1.46	.75
	A8	8	.6	4.5	2.50	1.54
	A9	3	1.6	4.7	2.76	1.70
	A10	38	.6	9.1	2.81	1.61
	A11	6	2.8	5.1	4.02	.86
vWa	A13	13	.7	6.2	2.14	1.69
	A14	14	.4	14.6	5.10	3.69
	A15	46	.65	11.66	6.46	2.34
	A16	56	.88	14.10	7.63	2.84
	A17	36	1.14	19.16	7.84	3.75
	A18	21	1.59	14.99	8.84	2.58
	A19	2	8.75	12.18	10.47	2.43

The distribution of alleles, the minimum, maximum, mean and standard deviation for percent stutter of each allele. Percent stutter was calculated as the ratio of peak heights of the stutter band divided by the allelic band.

The percent stutter was calculated by the computer software, ST*R call, during the analysis of the gel. The software determines the percent stutter by dividing the OD levels of the stutter band and the allele.

The average observed stutter for each STR locus was determine and is displayed in Fig. 4.

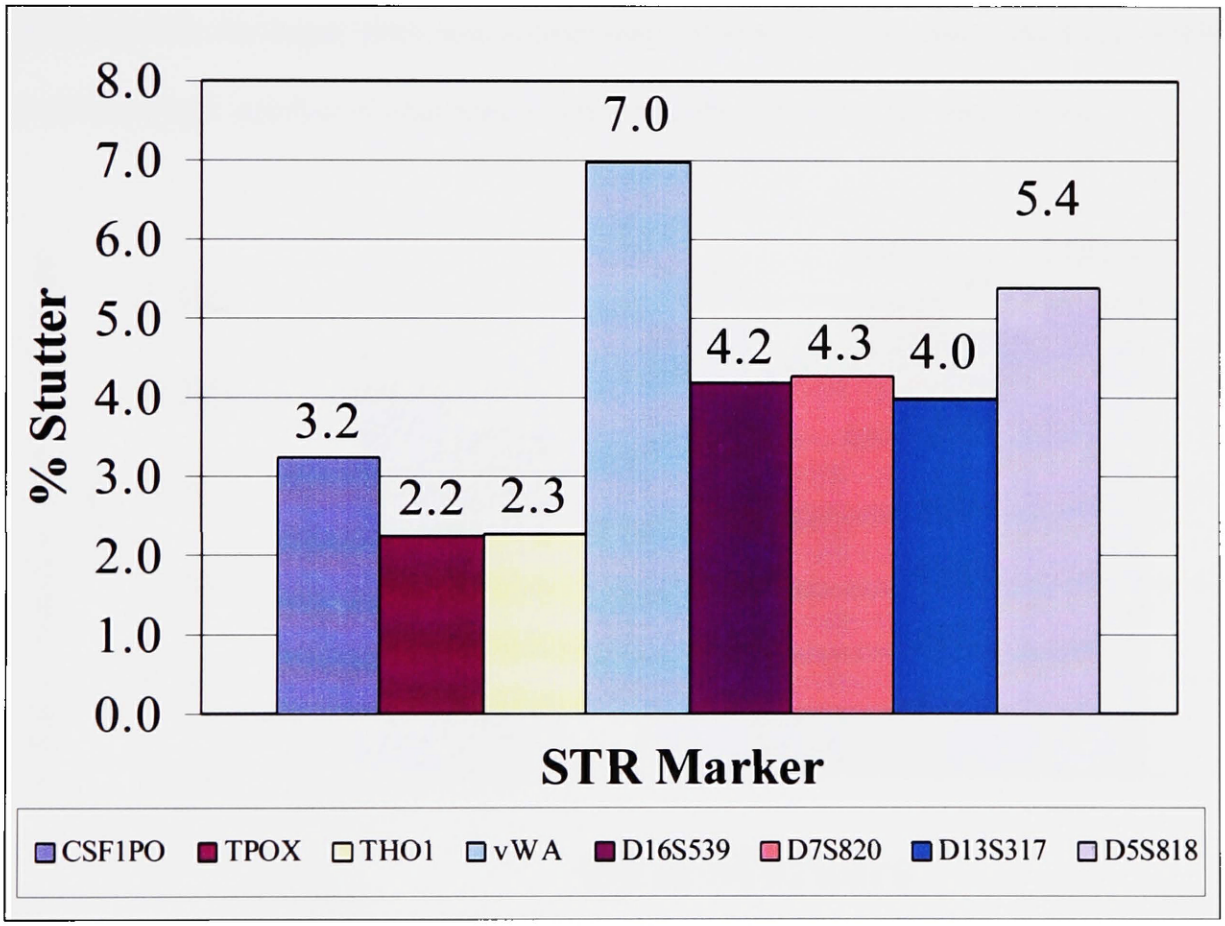


Fig. 4. The average percent stutter found in this study for each loci. The STR loci are along the X axis and the percent stutter is along the Y axis. The percent stutter was calculated by dividing the peak height of the stutter band by the peak height of the allelic band.

In forensic science, biological stains may be a mixture material from a heterozygote and heterozygote, homozygote and heterozygote, or homozygote and homozygote. If it

can be determined that in heterozygote samples, the stutter percentage of the larger allele is always smaller, larger or equal to the stutter percentage of the smaller allele, this information could assist in the determination of whether the unknown stain is a mixture. The heterozygous samples in this study were examined for the ratio of observed stutter between the smaller and larger alleles. Fig. 5 shows the ratios of stutter between the larger and smaller alleles in the heterozygous samples. In a majority of the samples the stutter level for the larger allele was greater than the smaller allele. Since the larger allele contains a larger number of core repeat units, this observation is not unexpected.

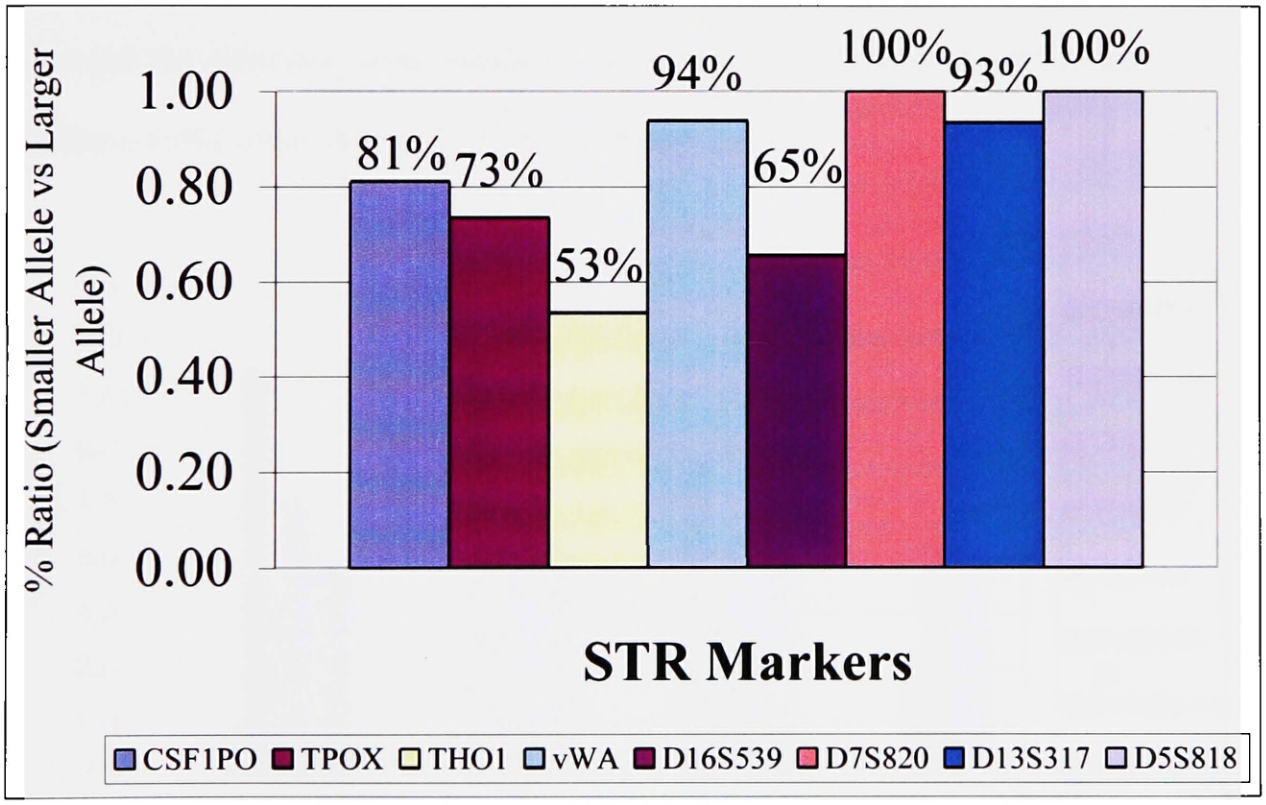


Fig. 5. Ratio of Heterozygous Stutter Percentages. The ratio of percent stutter for heterozygote samples was determined. The ratio was determined by dividing the stutter level of the smaller allele by the stutter level of the larger allele. Since the ratios of stutter changed for each locus, the ratio of larger/smaller percent stutter in questioned stains would not be useful.

Template Dilutions

Three samples were serially diluted to determine if the template concentration could affect the amount of observed stutter. The observed stutter did not change in a consistent fashion either between different samples or among different STR alleles. This could be due to the fact that the stutter production is not template concentration-dependent.

vWA showed a marked decrease in stutter levels as template concentration was decreased (data not shown). The other STR loci showed a consistent level of stutter regardless of the template concentration. Fig. 6 is an example of the relative stutter % values compared to template concentration for one sample. Analysis of this sample shows that the difference in the template concentration only accounts for 4% of the difference in the stutter levels (R^2=0.0434, data not shown).

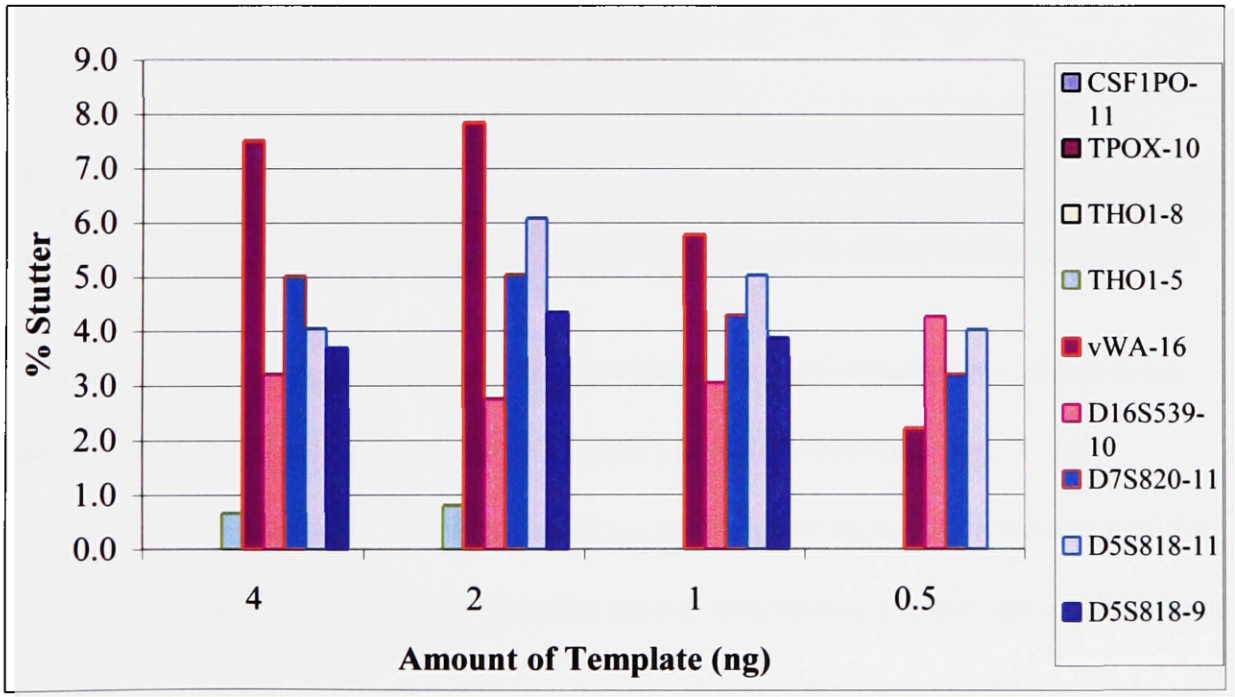


Fig. 6. Effect of Template Concentration on % Stutter Levels. The effect of decreasing the amount of template and the levels of stutter bands. The percent stutter was calculated by dividing the peak height of the stutter bands by the peak height of the allelic bands.

Consistency

Three different samples were amplified twelve times to determine if the level of stutter in a sample was consistent between amplifications. A visual inspection of the results from four alleles, as shown in Fig. 7 below, suggested that the levels of stutter were scattered and without an obvious pattern.

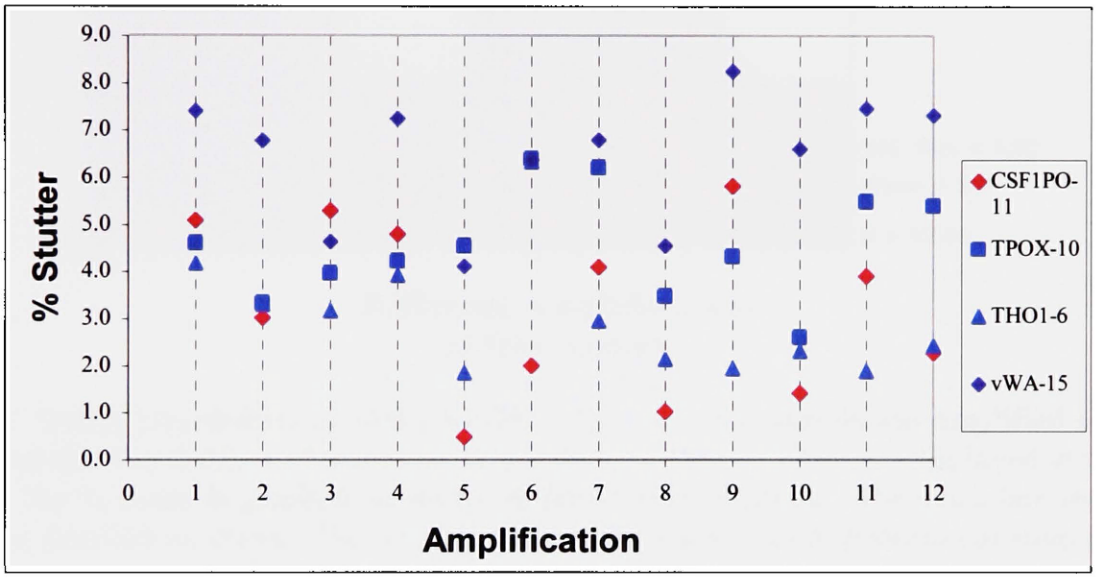


Fig. 7. Amount of CTTv Stutter in Repeated Amplifications in One Sample. The variability in CCTv stutter produced by twelve amplifications of one sample. The results of ANOVA analysis ($F_{1,46}$=.288, p=.594) indicated that repeated amplifications failed to produce a statistically significant difference in the amount of stutter between the alleles.

The distribution of the stutter levels did not show a normal distribution, which is not unexpected. Histograms of one allele from one sample are shown in Fig. 8. The distributions of stutter levels are compared to a normal distribution. Examination of the other alleles revealed a similar absence of a normal distribution pattern (data not shown).

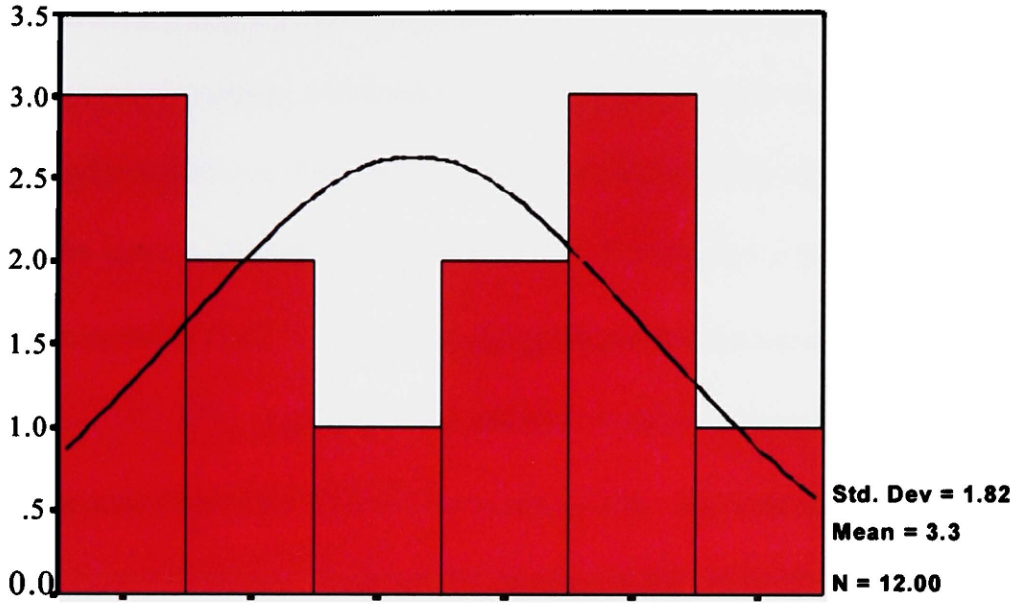


Fig. 8. Stutter Distribution of STR CSF1PO Allele 11. One sample was amplified six times and the distributions of stutter levels for the CSF1PO 11 allele are displayed in this graph. The % stutter is graphed verses the different amplifications. The black line shows a normal distribution curve. The samples did not show a normal distribution of stutter.

Thermal Cycle Conditions

The PCR process involves three distinct steps: (1) denaturation, separation of the DNA strands; (2) annealing, binding of the primers to specific sequences on the single-stranded DNA; and (3) extension, the enzymatic synthesis by the DNA polymerase of the complementary DNA strand initiating from the annealed primers. Each of the three steps have an optimal time and temperature. Using the manufacturer's recommended protocol as a starting point, the extension temperature, extension time, and overall temperature were varied to study the effects of stutter formation.

The first variation in the thermal cycle profile was to change the extension temperature. During the extension step the DNA polymerase synthesizes the second DNA strand by incorporating complementary nucleotides. The extension should start at the 3'-end of the annealed primers and stop at the end of the single-stranded DNA template.

Five samples were amplified with the PowerPlex™ multiplex system using the DNA polymerase Ampli*Taq* Gold™. The thermal cycle profile extension temperature was varied: 64°C, 67°C, 70°C, and 73°C. The amount of observed stutter was recorded as well as the amount of product (counts) from each allele. No temperature appears to be optimum for stutter reduction. None of the changes in the extension temperature eliminated stutter bands. In Fig. 9 the observed percent stutter is graphed against the extension temperature. The graph shows that the variation in extension temperature will only account for 0.13% of the percent stutter change (R^2=0.0013).

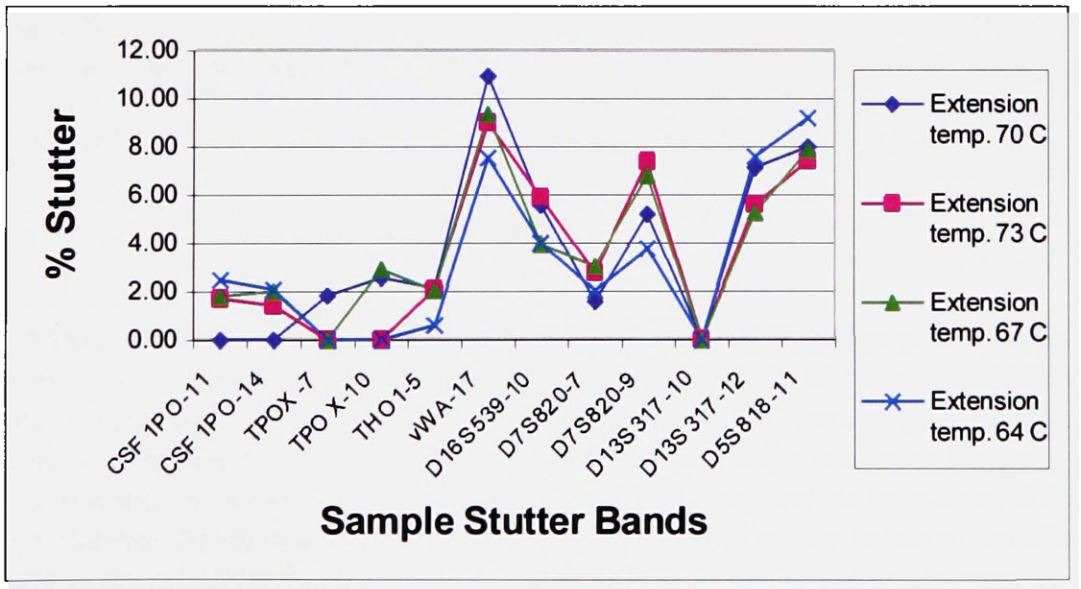


Fig. 9. Effect of Extension Temperature on Stutter. The % stutter produced in one sample where the extension temperature was varied above and below the recommended temperature of 70°C.

In the next experiment, the temperatures for each thermal cycle step were changed for five samples. The temperature of each step in the standard protocol was varied by +3°C or -3°C. Less stutter, as shown in Fig. 10, was produced by both changes; however, this decrease was accompanied by a decrease in the amount of amplified product. No amplified product was detected in many of the samples amplified with +3°C, as seen in Fig. 11.

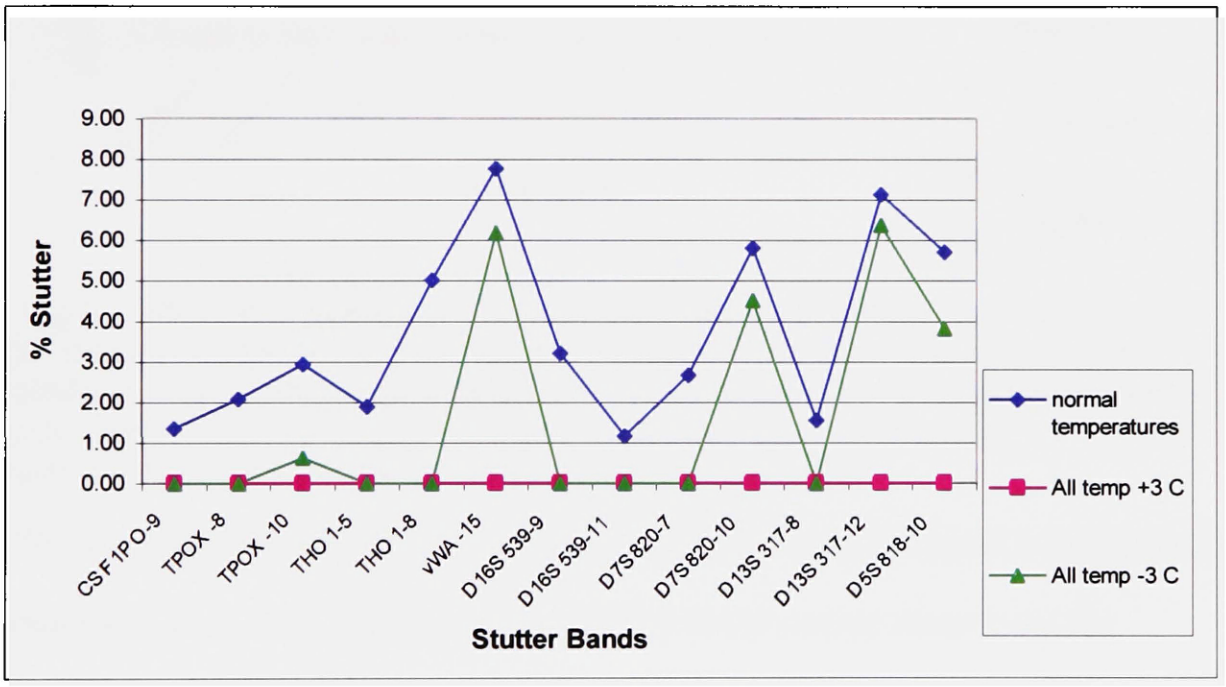


Fig. 10. Effect of Total Temperature on Stutter. The effect of % stutter produced by varying the total PCR amplification temperatures in one sample. In one amplification, the PCR temperatures were lowered by 3°C and then raised by 3°C for one amplification. The results of ANOVA analysis ($F_{1,70}$=.311, p=.082) on both the extension temperature variations and total temperature variations indicated that temperature variations failed to produce a statistically significant difference in the amount of stutter between the alleles. According to the adjusted R^2, 2.9% of the stutter change is explained by changes in the extension and thermal cycle temperatures.

It should be emphasized that the amount of amplified product can be very important in forensic samples. Fig. 11 shows the levels of amplified product for the DDDD alleles as

a function of temperature variation. A dramatic effect of raising all of the thermal cycle temperatures was the loss of amplified product in the DDDD alleles.

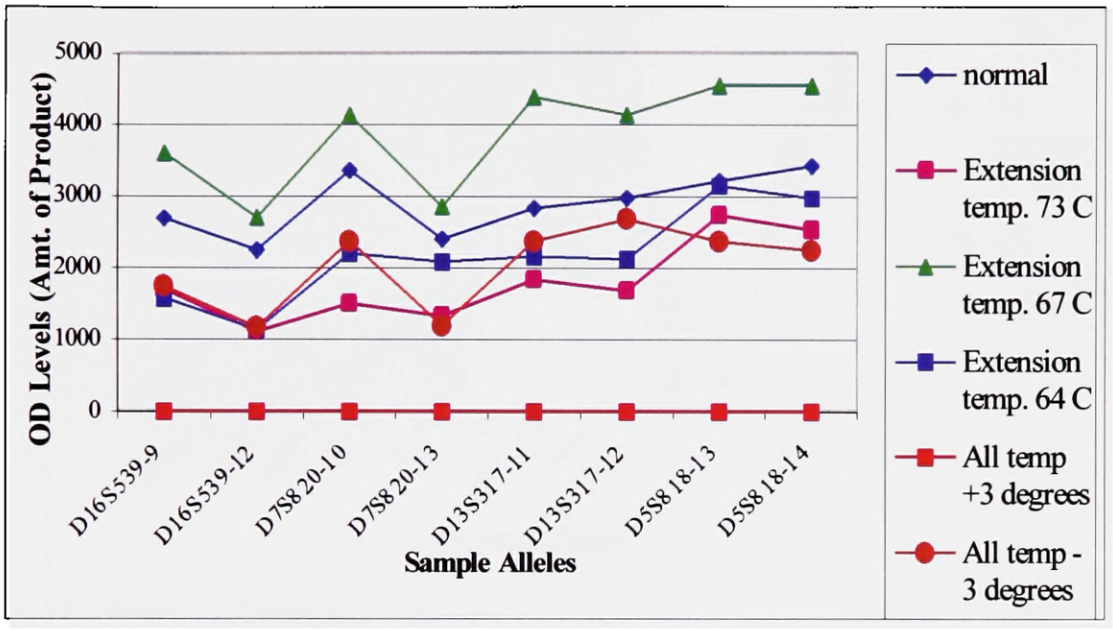


Fig. 11. Effect of Temperature Changes on the Amount of DDDD Amplified Product. The amount of amplified product for the DDDD alleles in one sample is plotted for variations in the thermal cycle temperatures. The greatest effect in variations of thermal cycle temperatures was seen by raising all the thermal cycle conditions by 3°C. Under these conditions, there was no production of amplified product.

The temperature extremes produced the least amount of amplified product and a general observation from all the experiments was that the lowest percent stutter is usually associated with the lowest level of amplified product.

Extension Time

The time for the extension step should allow the DNA polymerase to completely synthesis new daughter strands. To study the effects of extension time, the normal extension time of 1.5 minutes was either halved or doubled for five samples. The observed stutter was compared to the results from the recommended protocol.

Fig. 12 shows a comparison of the distribution of % stutter levels as a function of extension time.

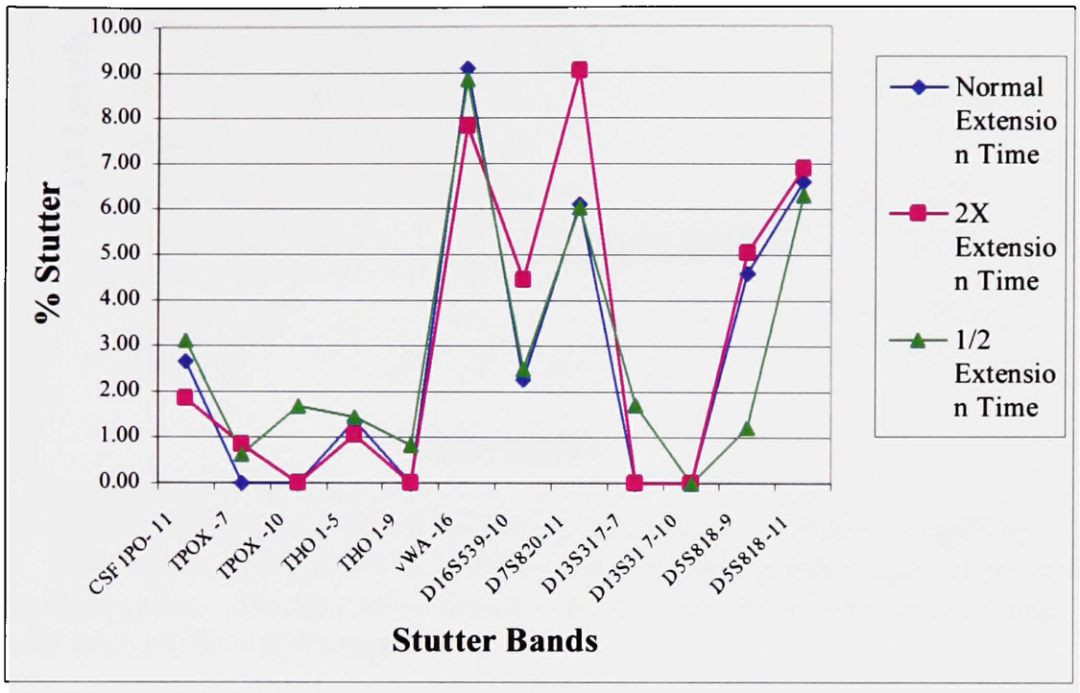


Fig. 12. Effect of Extension Time on % Stutter. The effect of variations in the extension time on % stutter for one sample. The results of ANOVA analysis ($F_{1,30}$=.370, p=.548) on both the extension temperature variations and total temperature variations indicated that temperature variations failed to produce a statistically significant difference in the amount of stutter between the alleles.

Fig. 13 graphs the amounts of amplified product against the variation in extension time for one sample. Altering the standard protocol lessened the amount of amplified product.

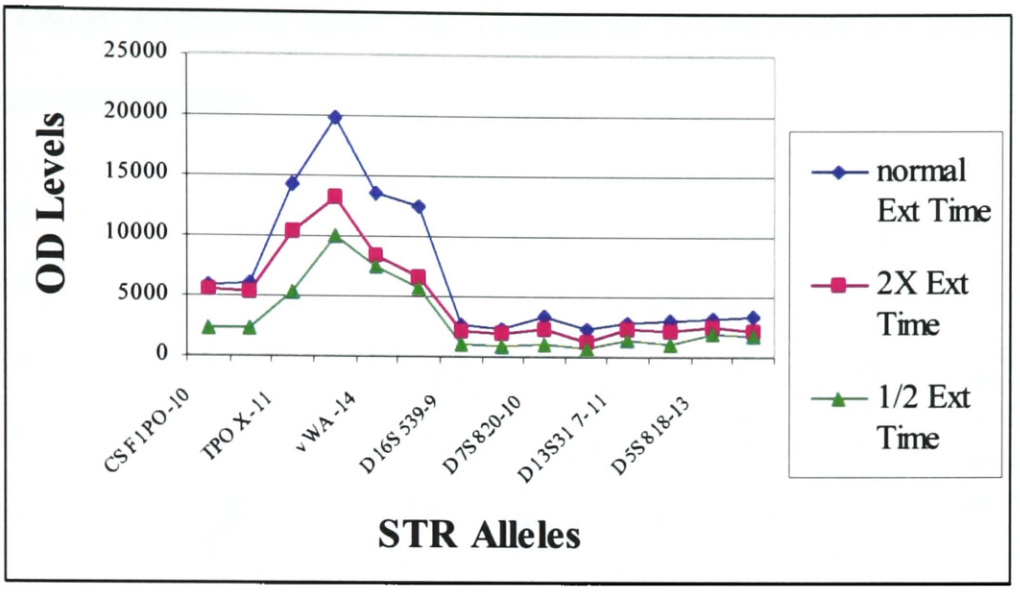


Fig. 13. Effect of Extension Time Changes on the Amount of DDDD Amplified Product. The amount of amplified product for one sample is graphed against variations in the extension time. Altering the extension time lowered the amount of amplified product for each allele in this sample.

A general observation from all the experiments was that the lowest percent stutter is usually associated with the lowest level of amplified product and less than 5% of the changes in the percent stutter can be attributed to changes in total temperature, extension temperature or extension time.

DNA Polymerase Concentration

Five different samples were amplified with various concentrations of Ampli*Taq* Gold™ DNA polymerase. The recommended concentration of Ampli*Taq* Gold™ DNA polymerase is 1 unit per reaction mixture. During this experiment, the amount of observed stutter was calculated with the amount of polymerase halved, doubled or quadrupled. I also measured the relative amount of amplified product for each allele, this is identified by the FMBIO as "counts". The samples with half of the recommended amount of DNA polymerase showed the least amount of amplified product and generally

the least amount of observed stutter. Fig. 14 shows the percent stutter as a function of polymerase concentration.

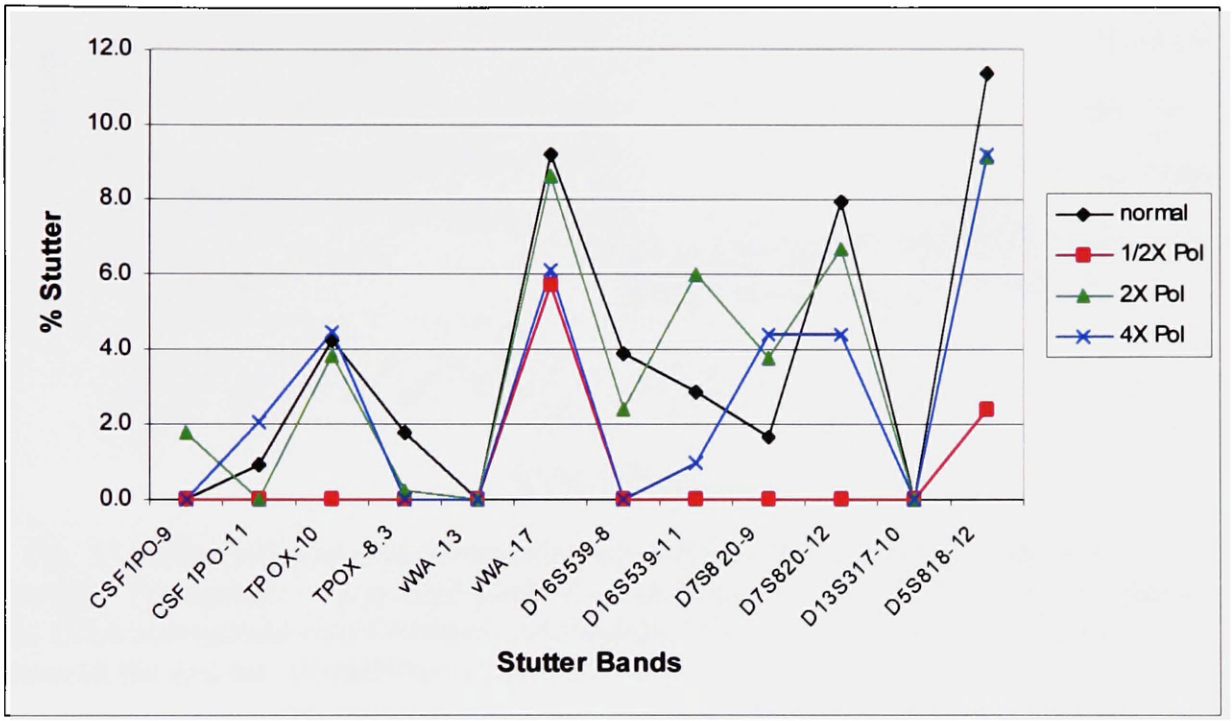


Fig. 14. Effect of DNA Polymerase Concentration on Stutter. The effect of variations in the DNA polymerase concentration on % stutter for one sample. The results of ANOVA analysis ($F_{1,47}$=1.107, p=.298) indicated that DNA polymerase variations failed to produce a statistically significant difference in the amount of stutter between the alleles. According to the adjusted R^2, 0.2 % of the stutter change is explained by changes in the DNA polymerase concentration.

Generally, the recommended amount of DNA polymerase produces the greatest amount of amplified product. Fig. 15 shows the difference in amount of amplified product as a function of DNA polymerase concentration for one sample. Altering the standard protocol lessened the amount of amplified product. A greater difference in the amplified products is seen in the four CTTv markers than in the four DDDD markers.

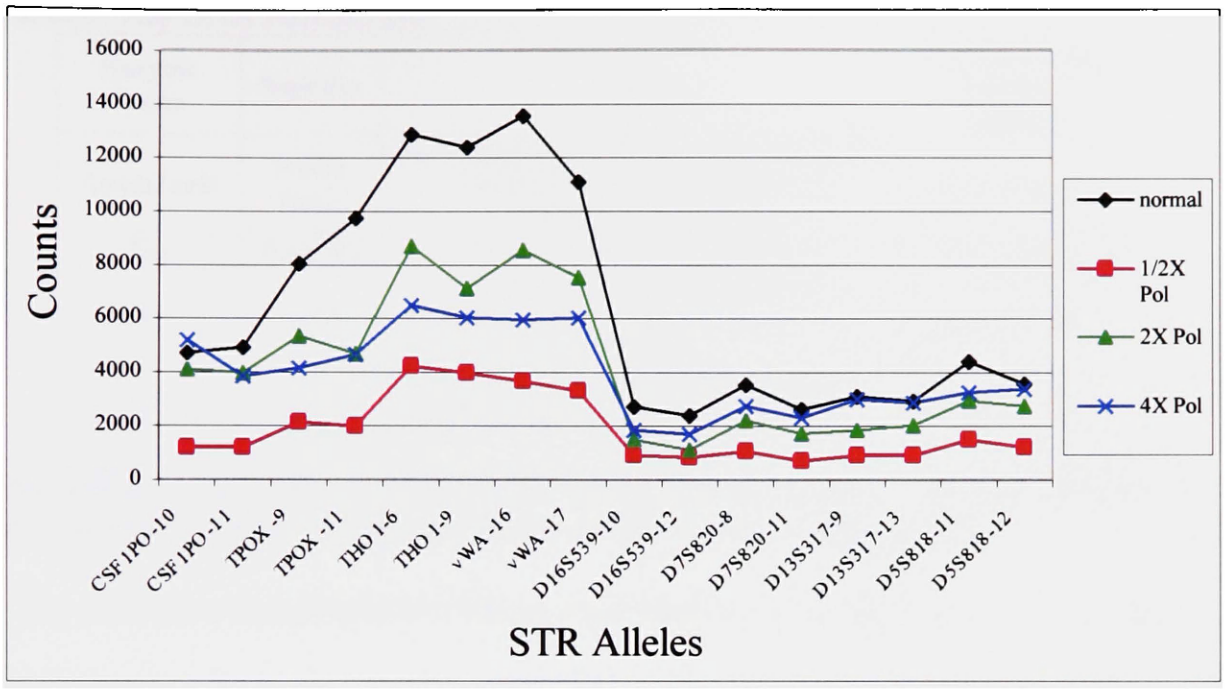


Fig. 15. Effect of DNA Polymerase Concentration on the Amount of Amplified Product. The amount of amplified product for one sample is graphed against variations in the DNA polymerase concentration. Altering the DNA polymerase concentration lowered the amount of amplified product for each allele in this sample.

This study with DNA polymerase concentration shows that amplifications with less DNA polymerase resulted in a decreased amount of stutter, but an unacceptable decrease in the amount of amplified product.

Polymerase Source

Four different *Taq* DNA polymerases from three different suppliers were used in the amplification of 12 replicate samples to determine if a difference in the amount of observed stutter could be identified. The different *Taq* DNA polymerases used in my study are listed in Table 8.

Table 8- *Taq* DNA Polymerases

Enzyme Name	Supplier	Source	Proofreading Enzyme present
Ampli*Taq*®	Perkin Elmer	Modified form of *Thermus aquaticus* DNA polymerase gene inserted into an *Escherichia coli* host.	None present
Taq	Promega	*Thermus aquaticus* strain YT1	None present
Taq	Sigma	Recombinant enzyme cloned from *Thermus aquaticus* and expressed in *E. coli*	None present
Red*Taq*™	Sigma	Recombinant enzyme cloned from *Thermus aquaticus* and expressed in *E. coli*	None present

The different *Taq* DNA polymerases that were used in my study. The table denotes the enzyme name, supplier, source and presence of any proofreading enzyme.

The difference between the two Sigma products is that Red*Taq*™ contains a red dye that *Taq* does not contain. As recommended by the supplier, the Sigma DNA polymerases used a reaction buffer of pH 8.3, (STR 10X buffer). The Promega and Perkin Elmer samples used a reaction buffer of pH 9.0, (GoldST*R buffer).

Fig. 16 shows the relationship between the different *Taq* DNA polymerases and the percent stutter for one sample. Ampli*Taq* has the highest and Red*Taq* has the lowest level of stutter, but these differences were not statistically significant.

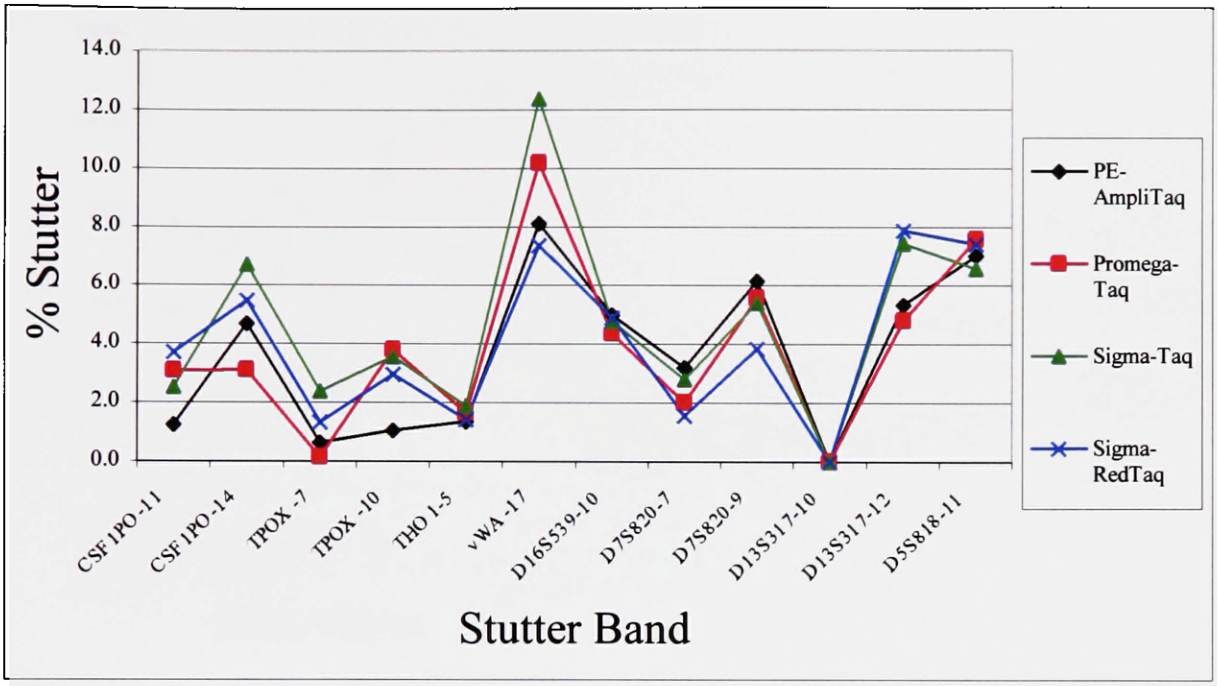


Fig. 16. Effect of *Taq* DNA Polymerase Source on % Stutter. The effect of variations in the *Taq* DNA polymerases on % stutter for one sample. The results of ANOVA analysis ($F_{3,181}$=2.358, p=1.457) indicated that variations in the *Taq* DNA polymerases failed to produce a statistically significant difference in the amount of stutter between the alleles.

For most of the samples, the *Taq* DNA polymerase from Perkin Elmer and Promega produced more amplified product. Fig. 17 demonstrates the difference in levels of amplified product produced by different *Taq* DNA polymerases from one sample. The relative differences in suppliers were more pronounced in the CSF1PO, TPOX, THO1 and vWA alleles. In four of the samples the differences in amounts from amplification by all of the suppliers was pronounced. In one sample, the *Taq* DNA polymerase from Promega produced more amplified product than *Taq* polymerase from the other suppliers. In another sample, the results were mixed (data not shown).

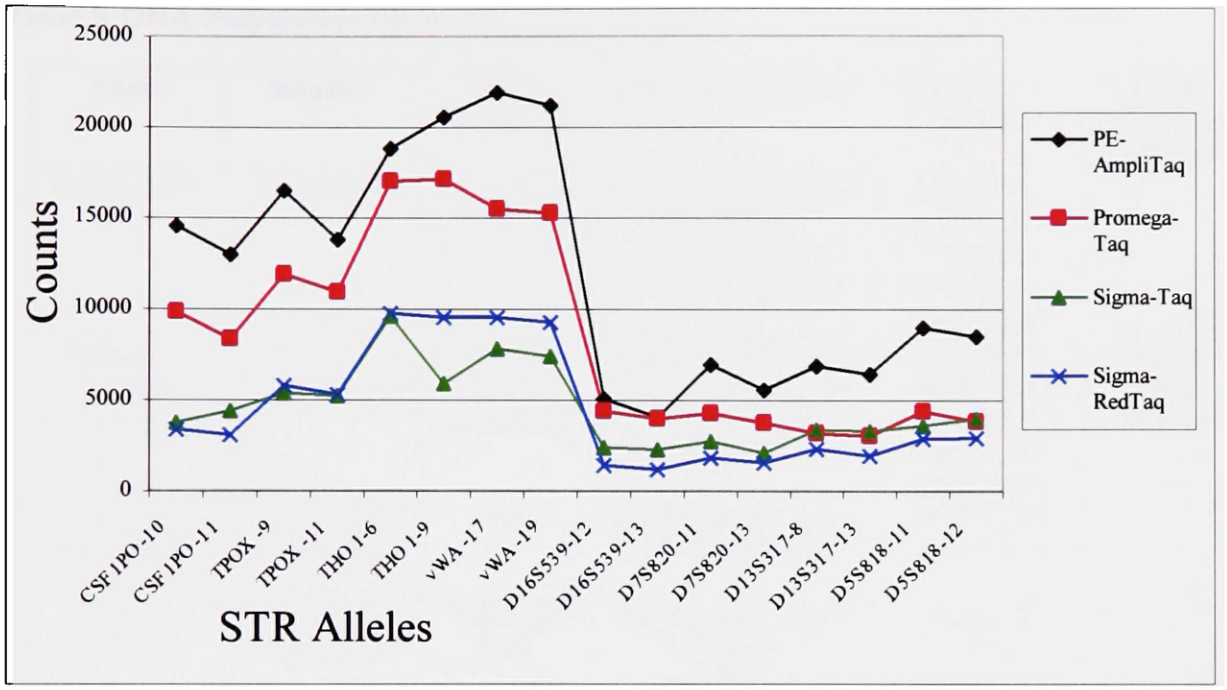


Fig. 17. Effect of *Taq* DNA Polymerase Source on the Amount of Amplified Product. The amount of amplified product for one sample is graphed against variations in *Taq* DNA polymerases. Altering the *Taq* DNA polymerase lowered the amount of amplified product for each allele in this sample.

In conclusion, the amount of amplified product and the amount of stutter are correlated, just as seen in the previous experiments. Red*Taq*™ produced the lowest amount of amplified product and stutter. Ampli*Taq*® produced the highest amount of both.

DNA Polymerase Mixtures

Taq DNA polymerase has little or no proofreading activity and thus a mixture of it with the addition of a thermostable DNA polymerase with a proofreading activity may increase the fidelity and reduce the amount of stutter.

Twelve samples were amplified with five different DNA polymerase mixtures. The amount of observed stutter was compared to the average observed stutter obtained from Ampli*Taq* Gold™. The different DNA polymerase mixtures used in my study are listed in Table 9.

Table 9-DNA Polymerase Mixtures

Name	Supplier	DNA Polymerase	Polymerase Source	Proofreading enzyme	Other substances present
DyNAyme™ EXT	Finnzymes	DyNAzymer II DNA Polymerase	Recombinant enzyme cloned from *T. brockianus* and expressed in *E. coli.*	Unspecified proofreading enzyme contained in the mixture	
Platinum™ *Taq* High Fidelity	Life Technologies	recombinant *Taq* DNA polymerase	Recombinant enzyme cloned from *Thermus aquaticus* and expressed in *E. coli*	Pyrococcus species GB-D thermostable polymerase-Deep Vent	Platinum *Taq* Antibody
Expand™ High Fidelity	Boehringer Mannheim	*Taq*	Information not given by manufacturer	*Pwo* DNA polymerase	
*Taq*Plus® Precision	Stratagene Corporation	*Taq* 2000 DNA Polymerase	Recombinant version of *Taq* DNA polymerase	Cloned *Pfu* DNA polymerase	
JumpStart™ *Taq*	Sigma Corporation	*Taq*	Recombinant enzyme cloned from *Thermus aquaticus* and expressed in E. coli	Deep Vent	*Taq* Start Antibody

The different *Taq* DNA polymerases that were used in my study, including the DNA polymerase, polymerase source, proofreading enzyme present, and the presence of any other substance.

For many of the samples, the DNA polymerase mixtures produced less stutter than *Taq* DNA polymerase alone. As seen in the other experiments, the DNA polymerase mixture that produced the lowest stutter also produced the least amount of amplified product. None of the DNA polymerase mixtures produced no stutter on all of the samples.

Fig. 18 displays the results for one sample of the percent stutter produced by the DNA polymerase mixtures. Amplifications with varied DNA polymerase mixtures failed to produce consistent levels of stutter.

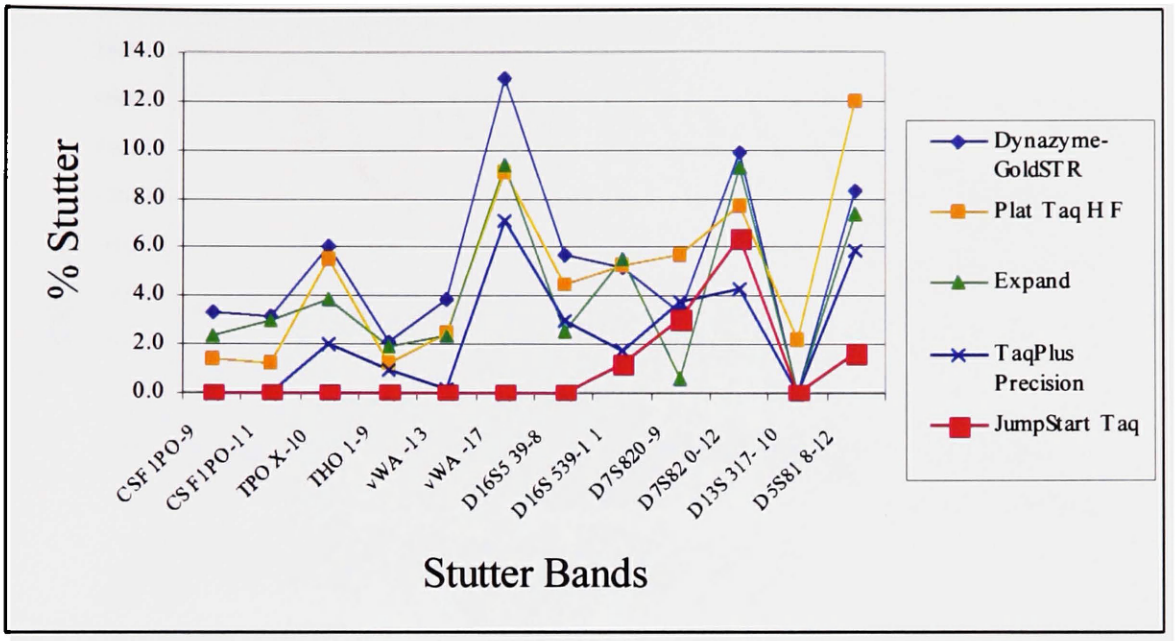


Fig. 18. Effect of DNA Polymerase Mixtures on Stutter. The effect of variations in the DNA polymerases on % stutter for one sample. In this sample, Jumpstart produced the least amount of stutter. Averaging all of the % stutter for these DNA polymerases showed that *Taq*Plus Precision produced the least amount of stutter (Table 13).

A sample of 238 stutter bands from four samples was compared using an analysis of variance test (data not shown). The test compared the stutter produced with a particular DNA mixture polymerase to the stutter produced by the other DNA polymerase mixtures. The results demonstrate a significant difference in stutter produced by different DNA polymerases, ($F_{4,234}$=3.781, P=.005).

Fig. 19 shows the amounts of amplified product as a function of the various DNA polymerase mixtures for one sample. For this sample, Platinum™ *Taq* High Fidelity consistently produced a higher level of amplified product. However, the production of higher levels of amplified product by this DNA polymerase mixture did not occur in all tested samples.

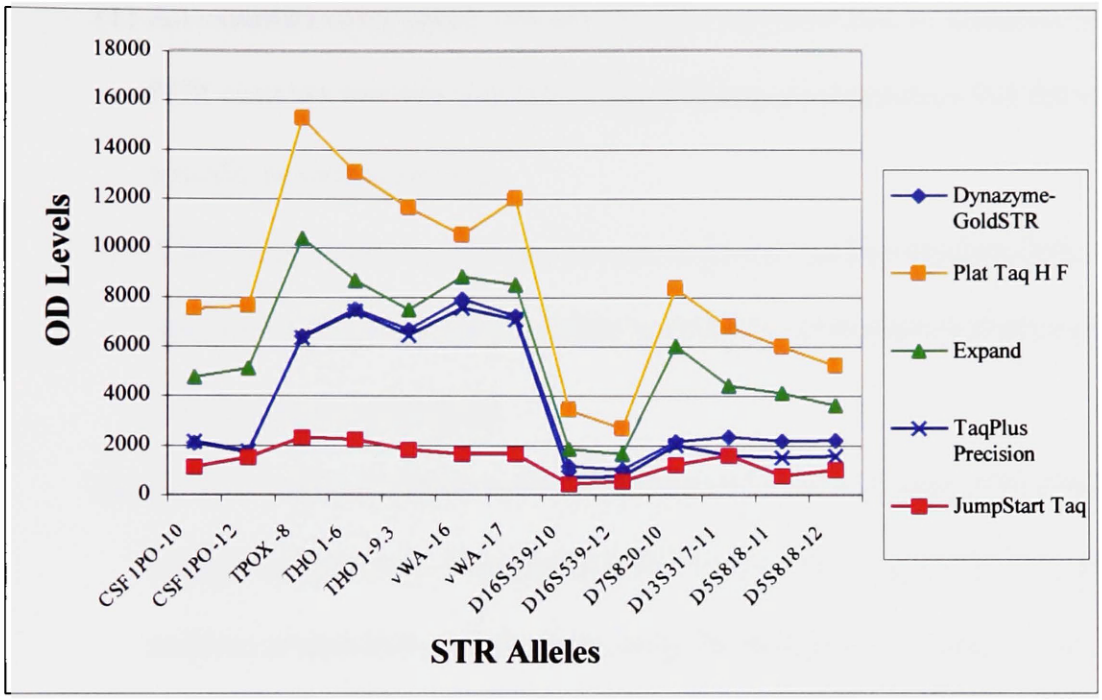


Fig. 19. Effect of DNA Polymerase Mixtures on the Amount of Amplified Product. The amount of amplified product for one sample is graphed against variations in DNA polymerase mixtures. Altering the DNA polymerase mixture altered the amount of amplified product for each allele in this sample.

JumpStart™ *Taq* did consistently produce a low amount of stutter, but, unfortunately it also consistently produced the lowest amount of amplified product (Table 13).

Again, less amplified product is correlated with less stutter, as seen in the previous experiments; however there is a statistically significant difference in stutter produced by the DNA polymerase mixtures.

Hot Start

PCR mixtures that are set up at ambient temperatures may allow nonspecific primer annealing and extension. These nonspecific primer products will compete with targeted sequences for dNTPs and primers and, as a result, reduce the yield of specific target products and create a confusing allele pattern. Several methods have been developed to eliminate the nonspecific primer products by various "hot start" protocols:

(1) An essential component, usually the DNA polymerase, is withheld from the PCR reaction mixture until the it has reached a temperature that favors specific primer annealing.

(2) A heat-labile wax or jelly barrier separates the reaction mixture from the template until the reaction mixture is heated to a temperature that favors specific primer annealing.

(3) Addition of an antibody to DNA *Taq* polymerase will inactive it until exposed to high temperatures. This method allows for PCR setup to be completed at ambient temperatures while preventing the formation of nonspecific primer products.

Twelve samples were amplified with three hot start *Taq* DNA polymerases from different suppliers. Table 10 lists the different DNA polymerases used in this experiment.

Table 10-Hot Start DNA Polymerases

Enzyme Name	Supplier	DNA Polymerase	Proofreading Enzyme	Polymerase Antibody
Ampli*Taq* Gold™	Perkin Elmer	*Taq*	none	Taq antibody
Platinum™ *Taq* High Fidelity	Life Technologies	*Taq*	None	Platinum Taq Antibody
JumpStart™ *Taq*	Sigma	*Taq*	Deep Vent	*Taq* Start Antibody

The "hot start" DNA polymerases used in this experiment, including the enzyme name, supplier, DNA polymerase, presence of an additional proofreading enzyme and polymerase antibody.

Fig. 20 shows the amount of stutter for one sample amplified with the hot start DNA polymerases.

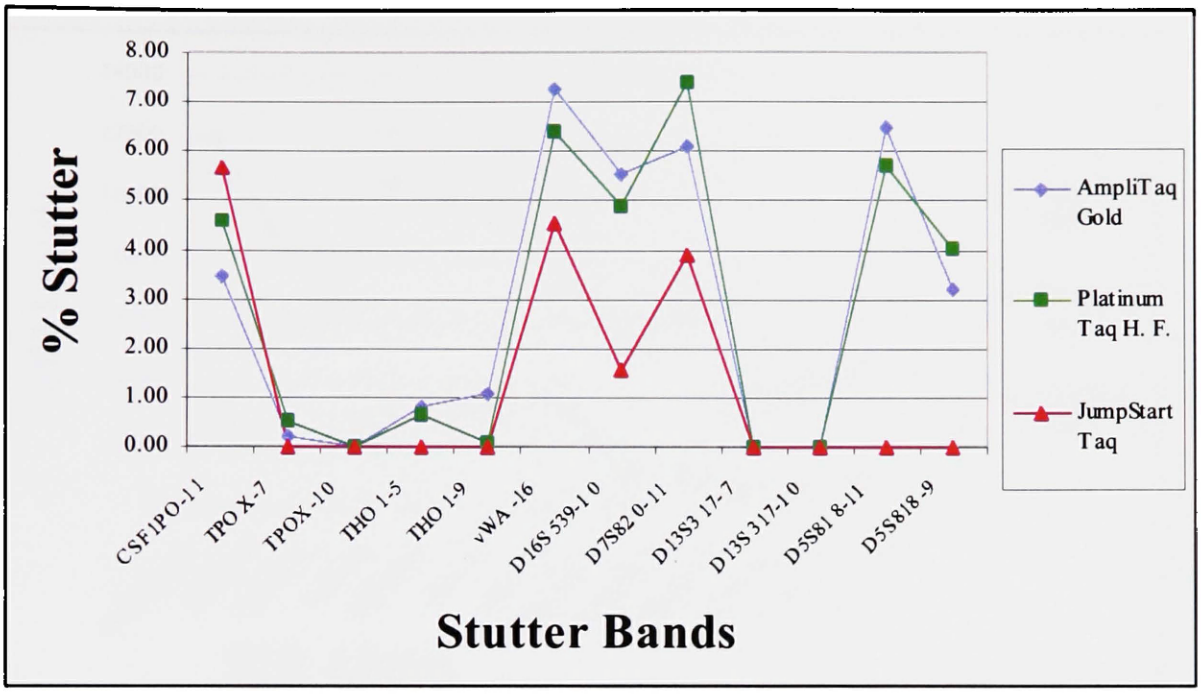


Fig. 20. Effect of Hot Start DNA Polymerases on Stutter. The effect of variations in the Hot Start DNA polymerases on % stutter for one sample. In this and all the samples in this experiment, Jumpstart™ produced the least amount of stutter (Table 13).

All 284 stutter bands were compared using an analysis of variance test (data not shown). The test compared the stutter produced with a particular hot start DNA polymerase to the stutter produced by the other hot start DNA polymerases. The results showed a significant statistical difference in stutter produced by different hot start DNA polymerases, ($F_{2,282}$=5.55; p=.004).

Fig. 21 demonstrates the difference in levels of amplified product produced by different *Taq* polymerases from one sample. There is an obvious difference in the amount of amplified product from each of the hot start DNA polymerases.

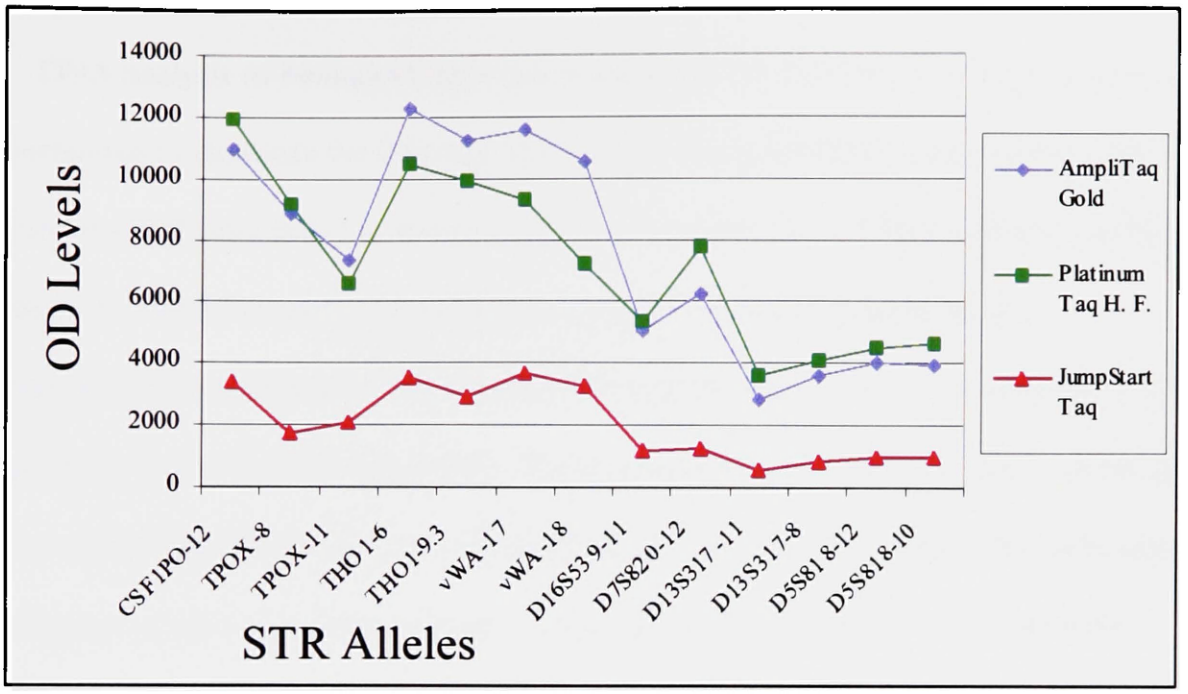


Fig. 21. Effect of Hot Start Polymerase on the Amount of Amplified Product. The amount of amplified product for one sample is graphed against variations in DNA Hot Start polymerases. Altering the DNA polymerase mixture altered the amount of amplified product for each allele in this sample. JumpStart™ produced both the least amount of stutter and amplified product.

Again, less amplified product is correlated with less stutter, just as seen in all the previous experiments. The amount of product is critical in forensic samples due to the limited amount of starting material. The choice of DNA polymerase must take into account the level of stutter (interpretation concerns) with the amount of amplified product (sufficient amounts of material for a result).

Chapter 4-Discussion

DNA analysis of biological materials using STRs has become a routine and powerful technique to eliminate the innocent and identify the donor of biological stains. One inherent difficulty found in forensic stains is the possibility of simple source mixtures, from the victim/suspect, suspect/suspect, suspect/unrelated individual, and victim/consensual partner. During the PCR amplification process, nontemplate DNA products can be formed with STRs. These unwanted DNA products bands usually appear one repeat unit below the true STR band and are called stutter bands, DNA polymerase slippage products, or shadow bands. The presence of these bands can confuse the interpretation of the DNA profile. My research investigated the PCR process and attempted to lower the amount of observed stutter by altering the concentrations of template and DNA polymerase, thermal cycler conditions, and type of DNA polymerase. None of the changes eliminated the production of stutter bands. However, stutter bands could be reduced, but this was always accompanied by a reduction in total amount of product.

The first portion of the research was to determine the levels of stutter found in the eight loci amplified in the PowerPlex™ system: CSF1PO, TPOX, THO1, vWA, D16S539, D7S820, D13S317, and D5S818. Published data has indicated that the ratio of the stutter to the allelic bands is less than 15% (Table 11) (Inman, 1997; Perkin Elmer Corporation, 1997). The observed levels of stutter in this study were less than the published stutter levels (Table 11). The locus with the largest percent stutter was vWA (Tables 5 and 9). This observation has been previously reported (Table 11) (Lins, 1998; Perkin Elmer Corporation, 1997; Promega Corporation, 1997; Riley, 1997).

Table 11-Stutter Levels

Locus	AmpF/STR Profiler Plus™, Published Stutter Levels	AmpF/STR COfiler ™, Published Stutter Levels	Promega, Published Stutter Levels	J. Streeter, Observed Stutter Averages
CSF1PO		10%	<5%	3%
TPOX		7%	<5%	2%
THO1		7%	<5%	2%
vWA	15%		>5%	7%
D16S539		15%	<5%	4%
D7S820	12%	12%	<5%	4%
D13S317	12%		<5%	4%
D5S818	12%		>5%	5%

The published levels of stutter and the levels of stutter of stutter found in my experiments. The AmpF/STR Profiler Plus™ (Perkin Elmer Corporation, 1997) and AmpF/STR COfiler ™ (Perkin Elmer Corporation, 1998) kits do not amplify the same STRs and therefore no stutter values were given for those loci. The Promega (Lins, 1998) kit amplifies all of the STRs.

Except at the THO1 locus, the percentage stutter increased with allele length at each locus (Table 7). All the loci studied in this experiment have alleles that differ by 4 base pairs, except THO1. The allele THO1 9.3 lacks an adenine in the seventh copy of the repeat (Puers, 1993). Analysis of stutter and sequences in the vWA marker, the STR with the highest amount of stutter (Walsh, 1996), shows that the vWA system has three repeat sequences. The TCTA sequence is repeated multiple times and represents the core sequence of the vWA repeat unit (Table 12). The number and location of the different repeat units vary between samples containing a typical verses low proportion of stutter. The proportion of stutter product relative to the main allele increases as the number of uninterrupted core repeat units increase (Bebenek, 1995; Kroutil, 1996). I observed that in each locus, with the exception of THO1, as the repeat number increased, the amount of observed stutter increased. Therefore stutter appears be mainly a result of the sequence context.

Table 12-PowerPlex™ Sequence Information

STR Locus	Allele(s)	Sequence structure 5'→3'	Sequence Variation 5'→3'
D16S539	5, 8-15	$[AGAT]_{5, 8-15}$	
D7S820	6-14	$[AGAT]_{6-14}$	
D13S317	7-15	$[TATC]_{7-15}$	
D13S317	10	$[TATC]_{10}[AATC]$	$[TATC]_{10}[TATC]$
D5S818	7-15	$[AGAT]_{7-15}$	
CSF1PO	6-15	$[AGAT]_{6-15}$	
TPOX	6-13	$[AATG]_{6-13}$	
THO1	5-11	$[AATG]_{5-11}$	
THO1	9.3	$[AATG]_6$[-ATG] $[AATG]_3$	
vWA	11,13	[TCTA] $[TCTG]_3[TCTA]_{7,9}$	
vWA	14	[TCTA] $[TCTG]_3[TCTA]_{10}$	[TCTA] $[TCTG]_5[TCTA]_3[TCCA]$ $[TCTA]_3$
vWA	15	[TCTA] $[TCTG]_4[TCTA]_{10}$	
vWA	16	[TCTA] $[TCTG]_4[TCTA]_{11}$	[TCTA] $[TCTG]_3[TCTA]_{12}$
vWA	17-20	[TCTA] $[TCTG]_4[TCTA]_{12-15}$	
vWA	21	[TCTA] $[TCTG]_6[TCTA]_{14}$	

The loci, alleles, sequence structure and sequence variation in the STR repeats. (Lins, 1998)

After determining the percent stutter at each locus, I examined the relationship between loci in heterozygous samples. Most of the heterozygous samples showed a larger amount of stutter in the larger of the two amplified alleles (Fig. 5). The same results were reported in the Perkin Elmer user manuals (Perkin Elmer Corporation, 1997; Perkin Elmer Corporation, 1998).

Replicate portions of several samples were amplified to determine the reproducibility of the observed stutter. The amount of observed stutter did vary within the samples (Fig. 7). Analysis of data showed that the difference in the sample only accounts for 2% difference in the stutter levels (R^2=.0207).

The next step was to vary the amount of template to determine if that had an effect on the quantity of observed stutter. While there was a general decrease in the amount of stutter with a decrease in template concentration (Fig. 6), the amount of template did not significantly affect the percent of observed stutter (R^2=.0434, data not shown). During the research, a general theme became evident, the amount of stutter correlated to the amount of amplified product. The conditions that produced the least amount of amplified product also produced the least percent of stutter.

The Promega Corporation kit contains standard protocols for preparation of the PCR master mix and thermal cycle conditions (Promega Corporation, 1997). The standardized protocol utilizes 1 unit of the DNA polymerase Ampli*Taq* Gold™ per sample reaction mixture. This amount was changed to determine the affect of DNA polymerase concentration. Generally, the recommended amount of DNA polymerase showed the greatest amount of amplified product. No clear-cut pattern could be discerned in comparing the stutter band production results from the different amounts of DNA polymerase (Fig. 14). The amounts of amplified product are more widely scattered in the CTTv marker set than the DDDD marker set (Fig. 15). It has been proposed that the use of additional DNA polymerase (Blake, 1992) could be used to overcome potential PCR inhibitors; this practice does not appear to have an affect on the amount of observed stutter (R^2=.0235).

The PCR process involves multiple repeated cycles of heating and cooling of the reaction mixtures for set time periods. During these cycles, millions of copies of selected areas of the DNA template are produced. Therefore, from very small stains, DNA typing results can be obtained. The DNA polymerase adds nucleotides onto the growing daughter strands during the extension step. Since the STRs are only 100-300 bases long, each tested DNA polymerase should have sufficient time to incorporate the appropriate nucleotides.

I varied the time and temperature of this step to study the affects of stutter production. None of the variations consistently lessened the amount of stutter production (Figs. 9 and 12). During the experiments, decreases in observed stutter were related to decreases in the total amount of amplified product (Figs. 11,15,17,19 and 21). The DNA polymerase molecule does not stay on the template DNA strand during the entire extension step. The DNA polymerase will dissociate and another binds. The number of nucleotides that the DNA polymerase will add until dissociation is called processivity. Processivity varies with the DNA Polymerase (Table 4).

PCR mixtures that are set up at ambient temperatures with the DNA polymerase *Taq* may allow nonspecific primer annealing and extension (Birch, 1996) and produce confusing results. The most common solution to this problem is to employ a "Hot Start" method. During "Hot Start," one essential component, usually the DNA polymerase, is withheld from the PCR reaction mixture until the solution has reached a temperature that favors specific primer annealing. One technique uses an antibody that will inactive the *Taq* DNA polymerase until exposed to high temperatures. This method allows for PCR setup to be completed at ambient temperatures while preventing the formation of

nonspecific primer products (Kellogg, 1994; Moretti, 1998). I compared several DNA polymerases that required a "Hot Start" for activation with regular *Taq* DNA polymerases. Analysis of the stutter bands did show a statistically significance difference in the level of stutter bands produced by the different "Hot-Start" DNA polymerases (data not shown). While AmpliTaq Gold™ did show a large variation in stutter, it also produced the greatest amount of amplified product (Fig. 21).

The identification and purification of *Thermus aquaticus* was reported in 1969 (Brock, 1969) and adapted for an automated PCR process in 1987 (Mullis, 1987). While *Taq* DNA polymerase lacks a 3'→5' proofreading exonuclease, it can catalyze highly accurate DNA synthesis *in vitro* (Eckert, 1990; Eckert, 1993; Saiki, 1986; Saiki, 1988) and is active over a broad range of temperatures (Innis, 1988; Lawyer, 1989). It has been suggested that exonucleolytic proofreading activity enhances fidelity (Kunkel, 1990). *Taq* has little or no proofreading activity and thus a mixture of *Taq* DNA polymerase with the addition of proofreading activity to form DNA polymerase mixtures may increase the fidelity (Cheng, 1994) and improve the performance of non-proofreading polymerases by correcting mismatches introduced during PCR (Cline, 1996). The error rate for a *Taq/Pfu* DNA polymerase mixture was found to be less than the error rate of *Taq* DNA polymerase (Cline, 1996). Expand™ High fidelity has been shown to have a three-fold increase in fidelity of DNA synthesis (8.5×10^{-6} error rate) when compared to *Taq* DNA polymerase (2.6×10^{-5} error rate) (Boehringer Mannheim, 1998).

The next portion of my research examined samples that were amplified with *Taq* DNA polymerase from different suppliers and then samples that were amplified with DNA proofreading polymerase mixtures. Four different *Taq* DNA polymerases from three

different suppliers were used in the amplification of 12 replicate samples to determine if a difference in the amount of observed stutter could be identified. Analysis of the data showed no significant statistical difference in the stutter levels between the different *Taq* polymerase samples (data not shown). Twelve samples were amplified with five different proofreading DNA polymerase mixtures. As seen in the other experiments, the DNA polymerase that produced the lowest stutter also produced the least amount of amplified product (Fig. 19). None of the DNA polymerase mixtures completely eliminated stutter. However, analysis did show a statistically significant difference in the levels of stutter produced by the different DNA polymerase mixtures, with JumpStart™ producing the lowest amount. The *Pfu* DNA polymerase was substituted in some experiments and it failed to provide any amplified product. *Pfu* DNA polymerase is isolated from the hyperthermophilic archaebacterium, *Pyrococcus furiosus*. The *Pfu* DNA polymerase was examined due to its eleven to twelve-fold greater replication fidelity than *Taq* DNA polymerase (Lundberg, 1991). Researchers have noted that *Pfu* may fail to amplify DNA on its own due to degradation of the PCR primers by its 3'-exonuclease activity (Barnes, 1994).

During this study I have examined the effect of the DNA polymerase on the production of STR stutter bands. STR stutter is a frameshift error that does not appear to be affected by the use DNA polymerases from different manufacturers, changes in the thermal cycle conditions or template and polymerase concentration. The production of stutter bands (frameshift errors) appears to be sequence-dependent. This theory was proposed previously (Kunkel, 1990; Kunkel, 1993).

The choice of a particular DNA polymerase, using a DNA polymerase mixture or a "Hot Start" DNA polymerase, can effect the level of stutter. One can reduce the level of stutter, but at the cost of reducing the level of amplified product, a key limiting factor in forensic science.

Ampli*Taq* Gold™ has shown in this research to be a good DNA polymerase for general use in the forensic laboratory. For most of the experiments, this DNA polymerase produced the most amplified product and acceptable levels of stutter. For samples that contain a sufficient amount of extracted DNA and a questionable level of stutter is observed with Ampli*Taq* Gold™, the use of another DNA polymerase should be considered. JumpStart™ *Taq* produced very low levels of stutter but low levels of amplified product. *Taq*Plus® Precision produced low levels of stutter and medium levels of amplified product (Table 13). Therefore each of the samples in that case could be reamplified with another DNA polymerase such as JumpStart™ *Taq* or *Taq*Plus® Precision to reduce the levels of stutter and assist in the interpretation of the DNA profile.

Another area of investigation that could assist in the interpretation of DNA profiles would be the implementation of STRs that contain a larger number of nucleotides in the repeat sequences. Recent work has shown that pentanucleotide repeats produce minimual levels of stutter bands (Schumm, 1998). Historically the amplified fragment length polymorphism D1S80, which contains 16 base pair repeats, did not produce stutter bands. One difficulty that the scientists could encounter with larger repeat sequences is the loss of intact DNA template in degraded samples or preferential amplification of smaller alleles. Validation tests with degraded samples, sample dilutions, and samples containing

mixtures of DNA should demonstrate if the pentanucleotide repeat STRs can be used in the forensic analysis of samples.

Table 13. Comparison of average stutter and average OD levels.

Name	Average Stutter	Average OD Levels
Platinum™ *Taq* High Fidelity	3.34	5413
Expand™	3.23	5756
Ampli*Taq* Gold™	3.02	5692
Sigma-*Taq*	2.77	3220
PE-Ampli*Taq*®	2.72	6847
Promega-*Taq*	2.56	6520
Dynazyme™	2.45	4052
LT-Platinum Taq	2.36	4935
Sigma-Red*Taq*	1.75	3020
JumpStart™ *Taq*	1.32	1949
*Taq*Plus® Precision	1.26	3373

Comparison of the average stutter and OD levels for all the DNA polymerases used in the experiments. The two DNA polymerases that produced the least amount of stutter were JumpStart™ and *Taq*Plus® Precision. However, there was a large difference in the average amount of amplified product.

Valuable information was gained during this study. While I was not able to eliminate stutter band production, I will be able to better interpret the DNA profiles knowing the limitations of the stutter bands.

Bibliography:

Abbotts, J., and Loeb, L., 1984, On the fidelity of DNA replication, *J. Biol. Chem.* **259**(11):6712-6714.

Barnes, W., 1994, PCR amplification of up to 35-kb DNA with high fidelity and high yield from λ bacteriophage templates, *Proc. Natl. Acad. Sci. USA* **91**:2216-2220.

Bebenek, K., Abbotts, J., Wilson, S., and Kunkel, T., 1993, Error-prone polymerization by HIV-1 reverse transcriptase, *J. Biol. Chem.* **268**(14):10324-10334.

Bebenek, K., Beard, W., Casas-Finet, J., Kim, H., Darden, T., Wilson, S., and Kunkel, T., 1995, Reduced frameshift fidelity and processivity of HIV-1 reverse transcriptase mutants containing alanine substitutions in helix H of the thumb subdomain, *J. Biol. Chem.* **270**(33):19516-19523.

Birch, D., 1996, Simplified hot start PCR, *Nature* **381**:445-446.

Blake, E., Mihalovich, J., Higuchi, R., Walsh, S. and Erlich, H., 1992, Polymerase chain reaction (PCR) amplification and human leukocyte antigen (HLA)-DQa oligonucleotide typing on biological evidence samples:casework experience, *J. Forensic Sciences* **37**(3):700-726.

Boehringer Mannheim, 1998, Expand high fidelity PCR system product insert.

Brock, T. and Freeze, H., 1969, *Thermus aquaticus* gen. n. and sp. n., a non-sporulating extreme thermophile, *J. Bacteriol.* **98 (No.1)**:289-297.

Browne, D., and Brock, A., 1954, Fingerprints Fifty Years of Scientific Crime Detection, E. P. Dutton & Co, p. 260.

Budowle, B., Moretti, T. , Keys, K. , Koons, B. and Smerick, J., 1997, Validation studies of the CTT STR multiplex system, *J. Forensic Sciences* **42**(No. 2):701-707.

Buel, E., 1998, Capillary electrophoresis STR analysis: comparison to gel-based systems, *J. Forensic Sciences* **43**(1):164-170.

Cheng, S., Fockler, C., Barnes, W. and Higuchi R., 1994, Effective amplification of long targets from cloned inserts and human genomic DNA, *Proc. Natl. Acad. Sci. USA* **91**:5695-5699.

Cline, J., Braman, J. and Hogrefe, H., 1996, PCR fidelity of *Pfu* DNA polymerase and other thermostable DNA polymerases, *Nucleic Acids Res.* **24**(18):3546-3551.

Comey, C., and Budowle, B., 1991, Validation studies on the analysis of the HLA DQa locus using the polymerase chain reaction, *J. Forensic Sciences* **36**(6):1633-1648.

Cosso, S., and Reynolds, R., 1995, Validation of the AmpliFLP D1S80 PCR amplification kit for forensic casework analysis according to TWGDAM guidelines, *J. Forensic Sciences* **40**(3):424-434.

Culliford, B. J., 1967, The determination of phosphoglucomutase (PGM) types in bloodstains, *J. Forensc Sci. Soc.* **7**:131.

Eckert, K., and Kunkel, T., 1990, High fidelity DNA synthesis by the *Thermus aquaticus* DNA polymerase, *Nucleic Acids Res.* **18**(13):3739-3744.

Eckert, K., and Kunkel, T., 1993, Effect of reaction pH on the fidelity and processivity of exonuclease-deficient klenow polymerase, *J. Biol. Chem.* **268**(18):13462-13471.

Edwards, A., Civitello, A., Hammond, H. A. and Caskey, C.T., 1991, DNA typing and genetic mapping with trimeric and tetrameric tandem repeats, *Am. J. Hum. Gen.* **49**:746-756.

Edwards, A., Hammond, H., Jin, L., Caskey, C.T. and Chakraborty R., 1992, Genetic variation at five trimeric and tetrameric tandem repeat loci in four human population groups, *Genomics* **12**:241-253.

Erlich, H., 1989, PCR technology (E. Henry, ed.), Stockton Press, pp. 246.

Erlich, H. A., Gelfland, D. and Sninsky, J.J., 1991, Recent advances in the polymerase chain reaction, *Science* **252**:643-651.

Fregeau, C., and Fourney, R., 1993, DNA typing with fluorescently tagged short tandem repeats: a sensitive and accurate approach to human identification, *BioTechniques* **15**(1):100-119.

Gibbs, R., 1990, DNA amplification by the polymerase chain reaction, *Anal. Chem.* **62**:1202-1214.

Hitachi, 1998, FMBIO user's manual, Hitachi Software.

Hite, J., Eckert, K., and Cheng, K., 1996, Factors affecting fidelity of DNA synthesis during PCR amplification of $d(C\text{-}A)_n.d(G\text{-}T)_n$ microsatellite repeats, *Nucleic Acids Res.* **24**(12):2429-2434.

Inman, K., and Rudin, N., 1997, An Introduction to Forensic DNA Analysis, CRC Press.

Innis, M., Myambo, K., Gelfand, D., and Brow, M., 1988, DNA sequencing with *Themus aquaticus* DNA polymerase and direct sequencing of polymerase chain reaction-amplified DNA, *Proc. Natl. Acad. Sci. USA* **85**:9436-9440.

Jeffreys, A., Wilson, V. and Thein, S.L., 1985, Individual-specific "fingerprints of human DNA", *Nature* **316**(4):76-79.

Jeffreys, A. J., Brookfield, J. F. Y. and Semeonoff, R., 1986, Positive identification of an immigration test case using human DNA fingerprints, *Nature* **317**:818-819.

Kellogg, D. E., Rybalkin, I., Chen, S., Mukhamedova, N., Vlasik, T., Siebert, P., and Chenchik, A., 1994, *Taq*start antibody: "hot start" PCR facilitated by a neutralizing monoclonal antibody directed against *Taq* DNA polymerase, *BioTechniques* **16**(6):1134-1137.

Keohavong, P. a. T., W., 1989, Fidelity of DNA polymerases in DNA amplification, *Proc. Natl. Acad. Sci. USA* **86**:9253-9257.

Kline, M., Duewer, D., Newall, P., Redman, J., Reeder, D. and Richard, M., 1997, Interlaboratory evaluation of short tandem repeat triplex CTT, *J. Forensic Sciences* **42**(5):897-906.

Kroutil, L., Register, K., Bebenek, K., and Kunkel, T., 1996, Exonucleolytic proofreading during replication of repetitive DNA, *Biochemistry.* **35**:1046-1053.

Kunkel, T., and Soni, A., 1988, Mutagenesis by transient misalignment, *J. Biol. Chem.* **263**(29):14784-14789.

Kunkel, T., 1990, Misalignment-mediated DNA synthesis errors, *Biochemistry* **29**(35):8003-8011.

Kunkel, T., 1992, DNA replication fidelity, *J. Biol. Chem.* **267**(26):18251-18254.

Kunkel, T., 1993, Slippery DNA and diseases, *Nature* **365**(16):207-208.

Landsteiner, K., 1901, Uber agglutination surscheinungen normalen menschlichen blutes, *Wien, and Klin. Worchenschr.* **14**:1132-1134.

Lawyer, F., Stoffel, S., Saiki, R., Myambo, K., Drummond, R., and Gelfand, D., 1989, Isolation, characteriaztion, and expression in *Escherichia coli* of the DNA polymerase gene from *Thermus aquaticus*, *J. Biol. Chem.* **264**(11):6427-6437.

Levinson, G., and Gutman, G., 1987, Slipped-strand mispairing: a major mechanism for DNA sequence evolution, *Mol. Biol. Evol.* **4**:203-221.

Lewin, B., 1997, Genes VI, Oxford University Press and Cell Press.

Lins, A., Sprecher, C., Puers, C. and Schumm, J., 1996, Multiplex sets for the amplification of polymorphic short tandem repeat loci-silver stain and fluorescence detection, *BioTechniques* **20**(5):882-889.

Lins, A., Micka, K., Sprecher, C., Taylor, J., Bacher, J., Rabbach, D., Bever, R., Creacy, S., and Schumm, J., 1998, Development and population study of an eight-locus short tandem repeat (STR) multiplex system, *J. Forensic Sciences* **43**(6):1-13.

Lundberg, K., Showmaker, D., Adams, M., Short, J., Sorge, J., and Mathur, E., 1991, High-fidelity amplification using a thermostable DNA polymerae isolated from *Pyrococcus furiosus*, *Gene* **108**:1-6.

Meldguard, M., and Morling, N., 1998, Detection and quantitative characterization of artificial extra peaks following polymerase chain reaction amplification of 14 short tandem repeat systems used in forensic investigations, *Electrophoresis* **18**:1928-1935.

Micka, K., Sprecher, C. , Lins, A. , Comey, C. , Koons, B. , Crouse, C. , Endean, D. Pirelli, K. , Lee, S. , Duda, N. and Schumm, J., 1996, Validation of multiplex polymorphic STR amplification sets developed for personal identification applications, *J. Forensic Sciences* **41**(No. 4):582-590.

Moretti, T., Koons, B., and Budowle, B., 1998, Enhancement of PCR amplification yield and specificity using ampli*Taq* gold DNA polymerase, *BioTechniques* **25**:716-722.

Mullis, K., and Faloona, F., 1987, Specific synthesis of DNA *in vitro* via a polymerase-catalyzed chain reaction, *Methods in Enzymology* **155**:335-350.

Murray, J., et al., 1994, A comprehensive human linkage map with centimorgan density, *Science* **265**:2049-2054.

Murray, V., Monchawin, C. and England, P. R., 1993, The determination of the sequences present in the shadow bands of a dinucleotide repeat PCR, *Nucleic Acids Res.* **21**(10):2395-2398.

Perkin Elmer Corporation, 1997, AmpFlSTR Profiler Plus user's manual.

Perkin Elmer Corporation, 1998, AmpFlSTR COfiler Plus user's manual, .

Promega Corporation, 1997, Geneprint, PowerPlex 1.1 System Technical Manual, Promega Corp.

Puers, C., Lins, A. , Sprecher, C. , Brinkmann, B. and Schumm, J., 1993, Analysis of polymorphic short tandem repeat loci using well-characterized allelic ladders, in: *Fourth International Symposium on Human Identification*, Scottsdale, AZ.

Reynolds, R., Sensabaugh, G. and Blake, E., 1991, Analysis of genetic markers in forensic DNA samples using the polymerase chain reaction, *Anal. Chem.* **63**(No. 1):2-15.

Ricciardone, M., Lins, A., Schumm, J. and Holland ,M., 1997, Multiplex systems for the amplification of short tandem repeat loci: evaluation of laser fluorescence detection, *BioTechniques* **23**(4):742-747.

Riley, G., Keam, V. ,Coleman, H. and Aulinskas, T., 1997, Validation of powerplex STR multiplex and amelogenin sex identification typing kits using the FMBIO fluorescent scanner: forensic casework and high-throughput convicted offender

databanking, in: *The Eighth International Symposium on Human Identification*, Scottsdale, AZ, pp. 53-55.

Saferstein, R., 1988, Forensic science handbook, Prentice Hall.

Saiki, R., Bugawan, T., Horn, G., Mullis, K., and Erlich, H., 1986, Analysis of enzymatically amplified *B*-globin and HLA-DQ*a* DNA with allele-specific oligonucleotide probes, *Nature* **324**(13):163-166.

Saiki, R., Gelfand, D., Stoffel, S., Scharf, S., Higuchi, R., Horn, G., Mullis, K., and Erlich, H., 1988, Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase, *Science* **239**:487-491.

Schlotterer, C., and Tautz, D., 1992, Slippage synthesis of simple sequence DNA, *Nucleic Acids Res.* **20**(No 2):211-215.

Schumm, J., Lins, A., Micka, K., Sprecher, C., Rabbach, D. and Bacher, J., 1997, Automated fluorescent detection of STR multiplexes-development of the *geneprint* powerplex and FFFL multiplexes for forensic and paternity applications, in: *First European Symposium on Human Identification.*

Schumm, J., 1998, Pentanucleotide repeats: highly polymorphic genetic markers displaying minimal stutter artifact., in: *The Ninth International Symposium on Human Identification*, Orlando Florida.

Sprecher, C., Puers, C. , Lins, A. and Schumm, J., 1996, General approach to analysis of polymorphic short tandem repeat loci, *BioTechniques* **20**(No. 2):266-276.

Thacker, J., Chalk , J., Ganesh, A. and North, P., 1992, A mechanism for deletion formation in DNA by human cell extracts: the involvement of short sequence repeats, *Nucleic Acids Res.* **20**(23):6183-6188.

Walsh, S., Fildes, N. and Reynolds, R., 1996, Sequence analysis and characterization of stutter products at the tetranucleotide repeat locus vWA, *Nucleic Acids Res.* **24**(14):2807-2812.

Worley, J., Ma, M. , Lee, S., Lins, A. , Schumm, J. and Mansfield, E., 1994, Rapid genetic typing on the fluorImager: human DNA quantitation, RFLP, D1S80, and short tandem repeat (STR) analysis, in: *Fifth International Symposium on Human Identification*, Scottsdale AZ, pp. 109-117.

Wraxall, B. a. C. B., 1968, A thin layer starch gel method for enzyme typing of bloodstains, *J. Forensic Sci. Soc.* **8**:81.

Wyman, A., and White, R., 1980, A highly polymorphic locus in human DNA, *Proc. Natl. Acad. Sci. USA* **77**(11):6754-6758.